FIGURE 1

CGGACGCGTGGGTGCGAGGCGAAGGTGACCGGGGACCGAGCATTTCAGATCTGCTCGGTAGA CCTGGTGCACCACCACCATGTTGGCTGCAAGGCTGGTGTCTCCGGACACTACCTTCTAGG GTTTTCCACCCAGCTTTCACCAAGGCCTCCCCTGTTGTGAAGAATTCCATCACGAAGAATCA ATGGCTGTTAACACCTAGCAGGGAATATGCCACCAAAACAAGAATTGGGATCCGGCGTGGGA GAACTGGCCAAGAACTCAAAGAGGCAGCATTGGAACCATCGATGGAAAAAATATTTAAAATT GATCAGATGGGAAGATGGTTTGTTGCTGGAGGGGGCTGCTGTTGGTCTTGGAGCATTGTGCTA CTATGGCTTGGGACTGTCTAATGAGATTGGAGCTATTGAAAAGGCTGTAATTTGGCCTCAGT ATGTCAAGGATAGAATTCATTCCACCTATATGTACTTAGCAGGGAGTATTGGTTTAACAGCT TTGTCTGCCATAGCAATCAGCAGAACGCCTGTTCTCATGAACTTCATGATGAGAGGCTCTTG GGTGACAATTGGTGTGACCTTTGCAGCCATGGTTGGAGCTGGAATGCTGGTACGATCAATAC CATATGACCAGAGCCCAGGCCCAAAGCATCTTGCTTGGTTGCTACATTCTGGTGTGATGGGT GCAGTGGTGGCTCCTCTGACAATATTAGGGGGTCCTCTTCTCATCAGAGCTGCATGGTACAC AGCTGGCATTGTGGGAGGCCTCTCCACTGTGGCCATGTGTGCGCCCAGTGAAAAGTTTCTGA ACATGGGTGCACCCCTGGGAGTGGGCCTGGGTCTCGTCTTTGTGTCCTCATTGGGATCTATG TTTCTTCCACCTACCACCGTGGCTGGTGCCACTCTTTACTCAGTGGCAATGTACGGTGGATT AGTTCTTTTCAGCATGTTCCTTCTGTATGATACCCAGAAAGTAATCAAGCGTGCAGAAGTAT CACCAATGTATGGAGTTCAAAAATATGATCCCATTAACTCGATGCTGAGTATCTACATGGAT **A**AGTGACTCAGCTTCTGGCTTCTCTGCTACATCAAATATCTTGTTTAATGGGGCAGATATGC ATTAAATAGTTTGTACAAGCAGCTTTCGTTGAAGTTTAGAAGATAAGAAACATGTCATCATA TTTAAATGTTCCGGTAATGTGATGCCTCAGGTCTGCCTTTTTTTCTGGAGAATAAATGCAGT AATCCTCTCCCAAATAAGCACACACATTTTCAATTCTCATGTTTGAGTGATTTTAAAATGTT TTGGTGAATGTGAAAACTAAAGTTTGTGTCATGAGAATGTAAGTCTTTTTTCTACTTTAAAA TTTAGTAGGTTCACTGAGTAACTAAAATTTAGCAAACCTGTGTTTGCATATTTTTTTGGAGT GCAGAATATTGTAATTAATGTCATAAGTGATTTGGAGCTTTGGTAAAGGGACCAGAGAGAAG GAGTCACCTGCAGTCTTTTGTTTTTTTAAATACTTAGAACTTAGCACTTGTGTTATTGATTA GCTGAACTTAACAAAACTGTTCATCCTGAAACAGGCACAGGTGATGCATTCTCCTGCTGTTG CTTCTCAGTGCTCTCTTTCCAATATAGATGTGGTCATGTTTGACTTGTACAGAATGTTAATC ATACAGAGAATCCTTGATGGAATTATATATGTGTGTTTTACTTTTGAATGTTACAAAAGGAA ATAACTTTAAAACTATTCTCAAGAGAAAATATTCAAAGCATGAAATATGTTGCTTTTTCCAG AATACAAACAGTATACTCATG

FIGURE 2

MLAARLVCLRTLPSRVFHPAFTKASPVVKNSITKNQWLLTPSREYATKTRIGIRRGRTGQEL
KEAALEPSMEKIFKIDQMGRWFVAGGAAVGLGALCYYGLGLSNEIGAIEKAVIWPQYVKDRI
HSTYMYLAGSIGLTALSAIAISRTPVLMNFMMRGSWVTIGVTFAAMVGAGMLVRSIPYDQSP
GPKHLAWLLHSGVMGAVVAPLTILGGPLLIRAAWYTAGIVGGLSTVAMCAPSEKFLNMGAPL
GVGLGLVFVSSLGSMFLPPTTVAGATLYSVAMYGGLVLFSMFLLYDTQKVIKRAEVSPMYGV
QKYDPINSMLSIYMDTLNIFMRVATMLATGGNRKK

FIGURE 3

GAAGGCTGCCTCGCTGGTCCGAATTCGGTGGCGCCACGTCCGCCCGTCTCCGCCTTCTGCAT GGTCGGCACGGGGGGTCGGGCGGTCTTGTGCATCTTGGCTACCTGTGGGTCGAAG<u>ATG</u>TCGG ACATCGGAGACTGGTTCAGGAGCATCCCGGCGATCACGCGCTATTGGTTCGCCGCCACCGTC GCCGTGCCCTTGGTCGGCAAACTCGGCCTCATCAGCCCGGCCTACCTCTTCCTCTGGCCCGA AGCCTTCCTTTATCGCTTTCAGATTTGGAGGCCAATCACTGCCACCTTTTATTTCCCTGTGG GTCCAGGAACTGGATTTCTTTATTTGGTCAATTTATATTTCTTATATCAGTATTCTACGCGA GATACCCAATGGGAGGAAGAAATTTTCTATCCACACCTCAGTTTTTGTACCGCTGG CTGCCCAGTAGGAGGAGGAGTATCAGGATTTGGTGTGCCCCCTGCTAGCATGAGGCGAGC TGCTGATCAGAATGGCGGAGGCGGGAGACACAACTGGGGCCAGGGCTTTCGACTTGGAGACC AGTGAAGGGCGGCCTCGGGGCAGCCGCTCTCTCTCTCCAGCCCACATTTCCTCCCAGTGCTGGGTG CACTTAACAACTGCGTTCTGGCTAACACTGTTGGACCTGACCCACACTGAATGTAGTCTTTC AGTACGAGACAAAGTTTCTTAAATCCCGAAGAAAAATATAAGTGTTCCACAAGTTTCACGAT TCTCATTCAAGTCCTTACTGCTGTGAAGAACAAATACCAACTGTGCAAAATTGCAAAACTGAC AAGACAGCCACGGATGAAGCGTTTCTCAGCTTTTGGAATTGCTTCGACTGACATCCGTTGTT TCAAGACTGTAGTGGAGTTGCAGCTAACATGGGTTAGGTTTAAACCATGGGGGATGCACCCC TTTGCGTTTCATATGTAGCCCTACTGGCTTTGTGTAGCTGGAGTAGTTGGGTTGCTTTGTGT TAGGAGGATCCAGATCATGTTGGCTACAGGGAGATGCTCTCTTTGAGAGGTCCTGGGCATTG TTTTGAGTCATGAATGTACAAGCTCTGTGAATCAGACCAGCTTAAATACCCACACCTTTTTT GTGTGTTATCCTAGGAAATGAGAGTTGGCAAACAACTTCTCATTTTGAATAGAGTTTGTGTG TACTTCTCCATATTTAATTTATATGATAAAATAGGTGGGGAGAGTCTGAACCTTAACTGTCA ATGAAGTTATTCCTCTTAAACATGGTTAGGAAGCTGATGACGTTATTGATTTTGTCTGGATT ATGTTTCTGGAATAATTTTACCAAAACAAGCTATTTGAGTTTTGACTTGACAAGGCAAAACA TGACAGTGGATTCTCTTTACAAATGGAAAAAAAAAATCCTTATTTTGTATAAAGGACTTCCC TTTTTGTAAACTAATCCTTTTTATTGGTAAAAATTGTAAATTAAAATGTGCAACTTG

FIGURE 4

MSDIGDWFRSIPAITRYWFAATVAVPLVGKLGLISPAYLFLWPEAFLYRFQIWRPITATFYF PVGPGTGFLYLVNLYFLYQYSTRLETGAFDGRPADYLFMLLFNWICIVITGLAMDMQLLMIP LIMSVLYVWAQLNRDMIVSFWFGTRFKACYLPWVILGFNYIIGGSVINELIGNLVGHLYFFL MFRYPMDLGGRNFLSTPQFLYRWLPSRRGGVSGFGVPPASMRRAADQNGGGGRHNWGQGFRL GDQ

Transmembrane domain:

amino acids 98-116, 152-172

N-myristoylation site.

amino acids 89-95, 168-174, 176-182, 215-221, 221-227, 237-243

Glycosaminoglycan attachment site.

amino acids 218-222

FIGURE 5

GGGGCCGCGGTCTAGGGCGGCTACGTGTTTTGCCATAGCGACCATTTTTGCATTAACTGGTTG GTAGCTTCTATCCTGGGGGCTGAGCGACTGCGGGCCAGCTCTTCCCCTACTCCCTCTCGGCT CCTTGTGGCCCAAAGGCCTAACCGGGGTCCGGCGGTCTGGCCTAGGGATCTTCCCCGTTGCC $\texttt{CCTTTGGGGCGGG} \underline{\textbf{ATG}} \texttt{GCTGCGGAAGAAGAAGACGAGGTGGAGTGGGTAGTGGAGAGCATCG}$ CGGGGTTCCTGCGAGGCCCAGACTGGTCCATCCCCATCTTGGACTTTGTGGAACAGAAATGT GAAGTTAACTGCAAAGGAGGGCATGTGATAACTCCAGGAAGCCCAGAGCCGGTGATTTTGGT GGCCTGTGTTCCCCCTTGTTTTTGATGATGAAGAAGCAAAATTGACCTATACAGAGATTC ATCAGGAATACAAAGAACTAGTTGAAAAGCTGTTAGAAGGTTACCTCAAAGAAATTGGAATT AATGAAGATCAATTTCAAGAAGCATGCACTTCTCCTCTTGCAAAGACCCATACATCACAGGC CATTTTGCAACCTGTGTTGGCAGCAGAAGATTTTACTATCTTTAAAGCAATGATGGTCCAGA AAAACATTGAAATGCAGCTGCAAGCCATTCGAATAATTCAAGAGAGAAATGGTGTATTACCT GACTGCTTAACCGATGGCTCTGATGTGGTCAGTGACCTTGAACACGAAGAGATGAAAATCCT GAGGGAAGTTCTTAGAAAATCAAAAGAGGAATATGACCAGGAAGAAGAAAGGAAGAAGAAAAA AACAGTTATCAGAGGCTAAAACAGAAGAGCCCACAGTGCATTCCAGTGAAGCTGCAATAATG CACAAAAAGGCCTGAAGATTCCTGGCTTAGAGCATGCGAGCATTGAAGGACCAATAGCAAAC TTATCAGTACTTGGAACAGAAGAACTTCGGCAACGAGAACACTATCTCAAGCAGAAGAGAGA TAAGTTGATGTCCATGAGAAAGGATATGAGGACTAAACAGATACAAAATATGGAGCAGAAAG GAAAACCCACTGGGGAGGTAGAGGAAATGACAGAGAAACCAGAAATGACAGCAGAGGAGAAA CAAACATTACTAAAGAGGAGATTGCTTGCAGAGAAACTCAAAGAAGAAGTTATTAATAAG**TA** CTTACACTG

FIGURE 6

MAAEEEDEVEWVVESIAGFLRGPDWSIPILDFVEQKCEVNCKGGHVITPGSPEPVILVACVP LVFDDEEESKLTYTEIHQEYKELVEKLLEGYLKEIGINEDQFQEACTSPLAKTHTSQAILQP VLAAEDFTIFKAMMVQKNIEMQLQAIRIIQERNGVLPDCLTDGSDVVSDLEHEEMKILREVL RKSKEEYDQEEERKRKKQLSEAKTEEPTVHSSEAAIMNNSQGDGEHFAHPPSEVKMHFANQS IEPLGRKVERSETSSLPQKGLKIPGLEHASIEGPIANLSVLGTEELRQREHYLKQKRDKLMS MRKDMRTKQIQNMEQKGKPTGEVEEMTEKPEMTAEEKQTLLKRRLLAEKLKEEVINK

N-glycosylation sites.

amino acids 224-228, 246-250, 285-289

N-myristoylation site.

amino acids 273-279

Amidation site.

amino acids 252-256

Cytosolic fatty-acid binding proteins.

amino acids 78-108

FIGURE 7

GGGCACAGCACATGTGAAGTTTTTGATGATGAAGAAGAAAGCAAATTGACCTATACAGAGAT
TCATCAGGAATACAAAGAACTAGTTGAAAAGCTGTTAGAAGGTTACCTCAAAGAAATTGGAA
TTAATGAAGATCAATTTCAAGAAGCATGCACTTCTCCTCTTGCAAAGACCCATACATCACAG
GCCATTTTTGCAACCTGTGTTGGCAGCAGAAGATTTTACTATCTTTAAAGCAATGATGGTCC
AGAAAAACATTGAAATGCAGCTGCAAGCCATTCGAATAATTCAAGAGAGAAAATGGTGTATTA
CCTGACTGCTTAACCGATGGCTCTGATGTGGTCAGTGACCTTGAACACGAAGAGATGAAAAT
CCTGAGGGAAGTTCTTAGAAAATCAAAAGAGGAATATGACCAGGAA

FIGURE 8

TAGCTTCTCCACGTATGGACCCTAAAGGCTACTGCTGCTACTACGGGGCTAGACAGTTACTG AGTGGA**ATG**GAAAAACAGTGCTGTAGTCATCCTGTAATATGCTCCTTGTCAACAATGTATAC ATTCCTGCTAGGTGCCATATTCATTGCTTTAAGCTCAAGTCGCATCTTACTAGTGAAGTATT CTGCCAATGAAGAAACAAGTATGATTATCTTCCAACTACTGTGAATGTGTGCTCAGAACTG AAATTTGAAATATGCTTCCTGGAAGGAATTCTCTGATTTCATGAAGTGGTCCATTCCTGCCT TTCTTTATTTCCTGGATAACTTGATTGTCTTCTATGTCCTGTCCTATCTTCAACCAGCCATG GCTGTTATCTTCTCAAATTTTAGCATTATAACAACAGCTCTTCTATTCAGGATAGTGCTGAA GAGGCGTCTAAACTGGATCCAGTGGGCTTCCCTCCTGACTTTATTTTTTGTCTATTGTGGCCT TGACTGCCGGGACTAAAACTTTACAGCACAACTTGGCAGGACGTGGATTTCATCACGATGCC TACAGCAAAGGAATGGACTTTTCCTGAAGCTAAATGGAACACCACAGCCAGAGTTTTCAGTC ACATCCGTCTTGGCATGGGCCATGTTCTTATTATAGTCCAGTGTTTTATTTCTTCAATGGCT AATATCTATAATGAAAAGATACTGAAGGAGGGGAACCAGCTCACTGAAAGCATCTTCATACA GAACAGCAAACTCTATTTCTTTGGCATTCTGTTTAATGGGCTGACTCTGGGCCTTCAGAGGA GTAACCGTGATCAGATTAAGAACTGTGGATTTTTTTTATGGCCACAGTGCATTTTCAGTAGCC CTTATTTTTGTAACTGCATTCCAGGGCCTTTCAGTGGCTTTCATTCTGAAGTTCCTGGATAA TCTTTGACTTCAGGCCCTCCCTGGAATTTTTCTTGGAAGCCCCATCAGTCCTTCTCTCTATA TTTATTATAATGCCAGCAAGCCTCAAGTTCCGGAATACGCACCTAGGCAAGAAAGGATCCG AGATCTAAGTGGCAATCTTTGGGAGCGTTCCAGTGGGGATGGAGAAGAACTAGAAAGACTTA CCAAACCCAAGAGTGATGAGTCAGATGAAGATACTTTC**TAA**CTGGTACCCACATAGTTTGCA GCTCTCTTGAACCTTATTTTCACATTTTCAGTGTTTTGTAATATTTATCTTTTCACTTTGATA AACCAGAAATGTTTCTAAATCCTAATATTCTTTGCATATATCTAGCTACTCCCTAAATGGTT CCATCCAAGGCTTAGAGTACCCAAAGGCTAAGAAATTCTAAAGAACTGATACAGGAGTAACA ATATGAAGAATTCATTAATATCTCAGTACTTGATAAATCAGAAAGTTATATGTGCAGATTAT TTTCCTTGGCCTTCAAGCTTCCAAAAAACTTGTAATAATCATGTTAGCTATAGCTTGTATAT ACACATAGAGATCAATTTGCCAAATATTCACAATCATGTAGTTCTAGTTTACATGCCAAAGT CTTCCCTTTTTAACATTATAAAAGCTAGGTTGTCTCTTGAATTTTGAGGCCCTAGAGATAGT CTGGCCATACCATAGATTTGGGATGATGTAGTCTGTGCTAAATATTTTTGCTGAAGAAGCAGT TTCTCAGACACAACATCTCAGAATTTTAATTTTTAGAAATTCATGGGAAATTGGATTTTTGT AATAATCTTTTGATGTTTTAAACATTGGTTCCCTAGTCACCATAGTTACCACTTGTATTTTA AGTCATTTAAACAAGCCACGGTGGGGCTTTTTTTCTCCTCAGTTTGAGGAGAAAAATCTTGAT AATTCAAGCTGTGACTATTGTATATCTTTCCAAGAGTTGAAATGCTGGCTTCAGAATCATAC CAGATTGTCAGTGAAGCTGATGCCTAGGAACTTTTAAAGGGATCCTTTCAAAAGGATCACTT AGCAAACACATGTTGACTTTTAACTGATGTATGAATATTAATACTCTAAAAATAGAAAGACC AGTAATATATAAGTCACTTTACAGTGCTACTTCACACTTAAAAGTGCATGGTATTTTTCATG GTATTTTGCATGCAGCCAGTTAACTCTCGTAGATAGAGAAGTCAGGTGATAGATGATATTAA AAATTAGCAAACAAAAGTGACTTGCTCAGGGTCATGCAGCTGGGTGATGATAGAAGAGTGGG CTTTAACTGGCAGGCCTGTATGTTTACAGACTACCATACTGTAAATATGAGCTTTATGGTGT CATTCTCAGAAACTTATACATTTCTGCTCTCCTTTCTCCTAAGTTTCATGCAGATGAATATA AGGTAATATACTATTATATAATTCATTTGTGATATCCACAATAATATGACTGGCAAGAATTG GTGGAAATTTGTAATTAAAATAATTATTAAACCT

FIGURE 9

MEKQCCSHPVICSLSTMYTFLLGAIFIALSSSRILLVKYSANEENKYDYLPTTVNVCSELVK
LVFCVLVSFCVIKKDHQSRNLKYASWKEFSDFMKWSIPAFLYFLDNLIVFYVLSYLQPAMAV
IFSNFSIITTALLFRIVLKRRLNWIQWASLLTLFLSIVALTAGTKTLQHNLAGRGFHHDAFF
SPSNSCLLFRSECPRKDNCTAKEWTFPEAKWNTTARVFSHIRLGMGHVLIIVQCFISSMANI
YNEKILKEGNQLTESIFIQNSKLYFFGILFNGLTLGLQRSNRDQIKNCGFFYGHSAFSVALI
FVTAFQGLSVAFILKFLDNMFHVLMAQVTTVIITTVSVLVFDFRPSLEFFLEAPSVLLSIFI
YNASKPQVPEYAPRQERIRDLSGNLWERSSGDGEELERLTKPKSDESDEDTF

Transmembrane domains:

amino acids 16-36 (type II), 50-74, 147-168, 229-250, 271-293, 298-318, 328-368

N-glycosylation sites.

amino acids 128-132, 204-208, 218-222, 374-378

Glycosaminoglycan attachment site.

amino acids 402-406

N-myristoylation sites.

amino acids 257-263, 275-281, 280-286, 284-290, 317-323

FIGURE 10

FIGURE 11

GTGGCTAAGGCTGCTACGAAGCGAGCTTGGGAGGAGCAGCGGCCTGCGGGGCAGAGGAGCAT CCCGTCTACCAGGTCCCAAGCGGCGTGGCCCGCGGGTCATGGCCAAAGGAGAAGGCGCCGAG AGCGGCTCCGCGGGGGGCTGCTACCCACCAGCATCCTCCAAAGCACTGAACGCCCGGCCCA GGTGAAGAAGAACCGAAAAAGAAGAACAACAGTTGTCTGTTTGCAACAAGCTTTGCTATG $\mathtt{CTATTGG}$ CTGGGATGCCATCACAGACCCCCTGGTGGGCCTCTGCATCAGCAAATCCCCCTGGACCTGCC TGGGTCGCCTTATGCCCTGGATCATCTTCTCCACGCCCCTGGCCGTCATTGCCTACTTCCTC ATCTGGTTCGTGCCCGACTTCCCACACGGCCAGACCTATTGGTACCTGCTTTTCTATTGCCT CTTTGAAACAATGGTCACGTGTTTCCATGTTCCCTACTCGGCTCTCACCATGTTCATCAGCA ACCGAGCAGACTGAGCGGGATTCTGCCACCGCCTATCGGATGACTGTGGAAGTGCTGGGCAC AGTGCTGGGCACGGCGATCCAGGGACAAATCGTGGGCCAAGCAGACACGCCTTGTTTCCAGG ACTTCAATAGCTCTACAGTAGCTTCACAAAGTGCCAACCATACACATGGCACCACTTCACAC AGGGAAACGCAAAAGGCATACCTGCTGGCAGCGGGGGTCATTGTCTGTATCTATATAATCTG TGCTGTCATCCTGATCCTGGGCGTGCGGGAGCAGAGAGACCCTATGAAGCCCAGCAGTCTG AGCCAATCGCCTACTTCCGGGGCCTACGGCTGGTCATGAGCCACGGCCCATACATCAAACTT ATTACTGGCTTCCTCTTCACCTCCTTGGCTTTCATGCTGGTGGAGGGGAACTTTGTCTTGTT TTGCACCTACACCTTGGGCTTCCGCAATGAATTCCAGAATCTACTCCTGGCCATCATGCTCT CGGCCACTTTAACCATTCCCATCTGGCAGTGGTTCTTGACCCGGTTTGGCAAGAAGACAGCT GTATATGTTGGGATCTCATCAGCAGTGCCATTTCTCATCTTGGTGGCCCTCATGGAGAGTAA CCTCATCATTACATATGCGGTAGCTGTGGCAGCTGGCATCAGTGTGGCAGCTGCCTTCTTAC TACCCTGGTCCATGCTGCCTGATGTCATTGACGACTTCCATCTGAAGCAGCCCCACTTCCAT GGAACCGAGCCCATCTTCTTCTCCTTCTATGTCTTCTCACCAAGTTTGCCTCTGGAGTGTC ACTGGGCATTTCTACCCTCAGTCTGGACTTTGCAGGGTACCAGACCCGTGGCTGCTCGCAGC CGGAACGTGTCAAGTTTACACTGAACATGCTCGTGACCATGGCTCCCATAGTTCTCATCCTG CTGGGCCTGCTCTTCAAAATGTACCCCATTGATGAGGAGAGGCGGCGGCAGAATAAGAA GGCCCTGCAGGCACTGAGGGACGAGGCCAGCAGCTCTGGCTGCTCAGAAACAGACTCCACAG AGCTGGCTAGCATCCTC**TAG**GGCCCGCCACGTTGCCCGAAGCCACCATGCAGAAGGCCACAG AAGGGATCAGGACCTGTCTGCCGGCTTGCTGAGCAGCTGGACTGCAGGTGCTAGGAAGGGAA CTGAAGACTCAAGGAGGTGGCCCAGGACACTTGCTGTGCTCACTGTGGGGCCCGGCTGCTCTG TGGCCTCCTGCCTCTGCCTGCCTGTGGGGCCAAGCCCTGGGGCTGCCACTGTGAATA TTAATGTTATTAATTTTCATAAAAGCTGGAAAGC

FIGURE 12

MWLRWALSLPPSSCLWAEPGMPSQTPWWASASANPPGPAWVALCPGSSSPRPWPSLPTSSSG SCPTSHTARPIGTCFSIASLKQWSRVSMFPTRLSPCSSATEQTERDSATAYRMTVEVLGTVL GTAIQGQIVGQADTPCFQDFNSSTVASQSANHTHGTTSHRETQKAYLLAAGVIVCIYIICAV ILILGVREQREPYEAQQSEPIAYFRGLRLVMSHGPYIKLITGFLFTSLAFMLVEGNFVLFCT YTLGFRNEFQNLLLAIMLSATLTIPIWQWFLTRFGKKTAVYVGISSAVPFLILVALMESNLI ITYAVAVAAGISVAAAFLLPWSMLPDVIDDFHLKQPHFHGTEPIFFSFYVFFTKFASGVSLG ISTLSLDFAGYQTRGCSQPERVKFTLNMLVTMAPIVLILLGLLLFKMYPIDEERRRQNKKAL QALRDEASSSGCSETDSTELASIL

FIGURE 13

GGGAAACGCAAAAGGCATACCTGCTGGCAGCGGGGGTCATTGTCTGTATCTATATAATCTGT
GCTGTCATCCTGATCCTGGGCGTGCGGGAGCAGAGAGAACCCTATGAAGCCCAGCAGTCTGA
GCCAATCGCCTACTTCCGGGGCCTACGGCTGGTCATGAGCCACGGCCCATACATCAAACTTA
TTACTGGCTTCCTCTTCACCTCCTTGGCTTTCATGCTGGTGGAGGGGAACTTTGTCTTGTTT
TGCACCTACACCTTGGGCTTCCGCAATGAATTCCAGAATCTACTCCTGGCCATCATGCTCTC
GGCCACTTTAACCATTCCCATCTGGCAGTGGTTCTTGACCCGGTTTGGCAAGAAGACAGCTG
TATATGTTGGGATCTCATCAGCAGTGCCATTTCTCATCTTGGTGGCCCTCATGGAGAGTAAC
CTCATCATTACATATGCGGTAGCTGTGGCAGCTGGCATCAGTGTGGCAGCTGCCTTCTTACT
ACCCTGGTCCATGCTGCCTGATGTCATTGACGACTTCCATCTGAAGCAGCCCCACTTCCATG
GAACCGAGCCCAT

FIGURE 14

GGGGCTTCGGCGCCAGCGCCAGCGCTAGTCGGTCTGGTAAGGATTTACAAAAGGTGCAGGT ATGAGCAGGTCTGAAGACTAACATTTTGTGAAGTTGTAAAACAGAAAACCTGTTAGAA**ATG**T GGTGGTTTCAGCAAGGCCTCAGTTTCCTTCCTTCAGCCCTTGTAATTTGGACATCTGCTGCT TTCATATTTTCATACATTACTGCAGTAACACTCCACCATATAGACCCGGCTTTACCTTATAT CAGTGACACTGGTACAGTAGCTCCAGAAAAATGCTTATTTGGGGCCAATGCTAAATATTGCGG CAGTTTTATGCATTGCTACCATTTATGTTCGTTATAAGCAAGTTCATGCTCTGAGTCCTGAA GAGAACGTTATCATCAAATTAAACAAGGCTGGCCTTGTACTTGGAATACTGAGTTGTTTAGG ACTTTCTATTGTGGCAAACTTCCAGAAAACAACCCTTTTTGCTGCACATGTAAGTGGAGCTG TGCTTACCTTTGGTATGGGCTCATTATATATGTTTGTTCAGACCATCCTTTCCTACCAAATG AGTAAGTGCACTTAGCATGCTGACTTGCTCATCAGTTTTTGCACAGTGGCAATTTTTGGGACTG ATTTAGAACAGAAACTCCATTGGAACCCCGAGGACAAAGGTTATGTGCTTCACATGATCACT ACTGCAGCAGAATGGTCTATGTCATTTTCCTTCTTTTGGTTTTTTCCTGACTTACATTCGTGA TTTTCAGAAAATTTCTTTACGGGTGGAAGCCAATTTACATGGATTAACCCTCTATGACACTG $\texttt{CACCTTGCCCTATTAACAATGAACGAACACGGCTACTTTCCAGAGATATT} \underline{\textbf{TGA}} \texttt{TGAAAGGAT}$ AAAATATTTCTGTAATGATTATGATTCTCAGGGATTGGGGAAAGGTTCACAGAAGTTGCTTA TTCTTCTGAAATTTTCAACCACTTAATCAAGGCTGACAGTAACACTGATGAATGCTGATA ATCAGGAAACATGAAAGAAGCCATTTGATAGATTATTCTAAAGGATATCATCAAGAAGACTA TTAAAAACACCTATGCCTATACTTTTTTATCTCAGAAAATAAAGTCAAAAGACTATG

FIGURE 15

MWWFQQGLSFLPSALVIWTSAAFIFSYITAVTLHHIDPALPYISDTGTVAPEKCLFGAMLNI
AAVLCIATIYVRYKQVHALSPEENVIIKLNKAGLVLGILSCLGLSIVANFQKTTLFAAHVSG
AVLTFGMGSLYMFVQTILSYQMQPKIHGKQVFWIRLLLVIWCGVSALSMLTCSSVLHSGNFG
TDLEQKLHWNPEDKGYVLHMITTAAEWSMSFSFFGFFLTYIRDFQKISLRVEANLHGLTLYD
TAPCPINNERTRLLSRDI

FIGURE 16

FIGURE 17

CCCACGCGTCCGCCGCCGCTGCGTCCCGGAGTGCAAGTGAGCTTCTCGGCTGCCCCGCGGG $\texttt{CCGGGGTGCGGAGCCGAC} \underline{\textbf{ATG}} \texttt{CGCCCGCTTCTCGGCCTCCTTCTGGTCTTCGCCGGCTGCAC}$ AGGCTGGAGGCAGGTCGCTGTGGTTCCCCTCCGACCTGGCAGAGCTGCGGGAGCTCTCTGAG GTCCTTCGAGAGTACCGGAAGGAGCACCAGGCCTACGTGTTCCTGCTCTTCTGCGGCGCCCTA CCTCTACAAACAGGGCTTTGCCATCCCCGGCTCCAGCTTCCTGAATGTTTTAGCTGGTGCCT TGCTACCTGCTCTCCAGTATTTTTGGCAAACAGTTGGTGGTGTCCTACTTTCCTGATAAAGT TGAGACTTTTCCCCATGACACCAAACTGGTTCTTGAACCTCTCGGCCCCAATTCTGAACATT CCCATCGTGCAGTTCTTCTCAGTTCTTATCGGTTTGATCCCATATAATTTCATCTGTGTGCAGACAGGGTCCATCCTGTCAACCCTAACCTCTCTGGATGCTCTTTTCTCCTGGGACACTG TCTTTAAGCTGTTGGCCATTGCCATGGTGGCATTAATTCCTGGAACCCTCATTAAAAAATTT AGTCAGAAACATCTGCAATTGAATGAAACAAGTACTGCTAATCATATACACAGTAGAAAAAGA CACA**TGA**TCTGGATTTTCTGTTTGCCACATCCCTGGACTCAGTTGCTTATTTGTGTAATGGA TGTGGTCCTCTAAAGCCCCTCATTGTTTTTGATTGCCTTCTATAGGTGATGTGGACACTGTG CATCAATGTGCAGTGTCTTTTCAGAAAGGACACTCTGCTCTTGAAGGTGTATTACATCAGGT TTTCAAACCAGCCCTGGTGTAGCAGACACTGCAACAGATGCCTCCTAGAAAATGCTGTTTGT GGCCGGGCGCGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCCGGTGATTC ACAAGGTCAGGAGTTCAAGACCAGCCTGGCCAAGATGGTGAAATCCTGTCTCTAATAAAAAT ACAAAATTAGCCAGGCGTGGTGGCAGGCACCTGTAATCCCAGCTACTCGGGAGGCTGAGGC AGGAGAATTGCTTGAACCAAGGTGGCAGAGGTTGCAGTAAGCCAAGATCACACCACTGCACT CCAGCCTGGGTGATAGAGTGAGACACTGTCTTGAC

FIGURE 18

MRPLLGLLLVFAGCTFALYLLSTRLPRGRRLGSTEEAGGRSLWFPSDLAELRELSEVLREYR
KEHQAYVFLLFCGAYLYKQGFAIPGSSFLNVLAGALFGPWLGLLLCCVLTSVGATCCYLLSS
IFGKQLVVSYFPDKVALLQRKVEENRNSLFFFLLFLRLFPMTPNWFLNLSAPILNIPIVQFF
FSVLIGLIPYNFICVQTGSILSTLTSLDALFSWDTVFKLLAIAMVALIPGTLIKKFSQKHLQ
LNETSTANHIHSRKDT

Important features:

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 101-123, 189-211

N-glycosylation sites.

amino acids 172-176, 250-254

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 240-244, 261-265

N-myristoylation site.

amino acids 13-19, 104-110, 115-121, 204-210

Amidation site.

amino acids 27-31

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 4-15

Protein splicing proteins.

amino acids 25-31

Sugar transport proteins.

amino acids 162-172

FIGURE 19

CCGAGGCGGGAGGCCCGAGGGGGGCCCCGCATGAATCATTGTAGTCAATCATTTT CCAGTTCTCAGCCGCTCAGTTGTGATCAAGGGACACGTGGTTTTCCGAACTGCCAGCTCAGAA TAGGAAAATAACTTGGGATTTTATATTGGAAGAC**ATG**GATCTTGCTGCCAACGAGATCAGCA TCAGAGAAGGCAATTGAAAAATTTATCAGACAGCTGCTGGAAAAGAATGAACCTCAGAGACC CCCCCGCAGTATCCTCTCTTATAGTTGTGTATAAGGTTCTCGCAACCTTGGGATTAATCT TGCTCACTGCCTACTTTGTGATTCAACCTTTCAGCCCATTAGCACCTGAGCCAGTGCTTTCT GGAGCTCACACCTGGCGCTCACTCATCACATTAGGCTGATGTCCTTGCCCATTGCCAA CAGACTTTGACCCCTGGTGGACAAACGACTGTGAGCAGAATGAGTCAGAGCCCATTCCTGCC AACTGCACTGGCTGTGCCCAGAAACACCTGAAGGTGATGCTCCTGGAAGACGCCCCAAGGAA ATTTGAGAGGCTCCATCCACTGGTGATCAAGACGGGAAAGCCCCTGTTGGAGGAAGAGATTC AGCATTTTTTGTGCCAGTACCCTGAGGCGACAGAAGGCTTCTCTGAAGGGTTTTTCGCCAAG TGGTGGCGCTGCTTTCCTGAGCGGTGGTTCCCATTTCCTTATCCATGGAGGAGACCTCTGAA CCTCTTTAAACAAGTGCTCCTTTCTTCACCCAGAACCTGTTGTGGGGAGTAAGATGCATAAG GTGCCGAAGACATTGTCAGTCTGTGGCCATGCCAATAGAGCCAGGGGATATCGGCTATGTCG ACACCACCACTGGAAGGTCTACGTTATAGCCAGAGGGGTCCAGCCTTTGGTCATCTGCGAT GGAACCGCTTTCTCAGAACTG**TAG**GAAATAGAACTGTGCACAGGAACAGCTTCCAGAGCCGA AAACCAGGTTGAAAGGGGAAAAATAAAAACAAAAACGATGAAACTGCAAAAA

FIGURE 20

MDLAANEISIYDKLSETVDLVRQTGHQCGMSEKAIEKFIRQLLEKNEPQRPPPQYPLLIVVY
KVLATLGLILLTAYFVIQPFSPLAPEPVLSGAHTWRSLIHHIRLMSLPIAKKYMSENKGVPL
HGGDEDRPFPDFDPWWTNDCEQNESEPIPANCTGCAQKHLKVMLLEDAPRKFERLHPLVIKT
GKPLLEEEIQHFLCQYPEATEGFSEGFFAKWWRCFPERWFPFPYPWRRPLNRSQMLRELFPV
FTHLPFPKDASLNKCSFLHPEPVVGSKMHKMPDLFIIGSGEAMLQLIPPFQCRRHCQSVAMP
IEPGDIGYVDTTHWKVYVIARGVQPLVICDGTAFSEL

FIGURE 21

FIGURE 22

CCCACGCGTCCGCCCACGCGTCCGGCTGAACACCTCTTCTTTGGAGTCAGCCACTGATGAGG CACCGAATGCGCACTTCATCGACAAACAGGTACAGCCAACCATGCCCAGTTCGAAATGGACACGTATGCTAAGAGCCACGACCTTATGTCAGGTTTCTGGAATGCCTGCTATGACATGCTTATGAGCAGTGGGCAGCGCCCAGTGGGAGCCCCAGAGTCGTCGGGCCTTCCAGGAGC TATGAGCAGTGGGCAGCGCCCAGTGGGGAGCGCCCCAGAGTCGTCGGGCCTTCCAGGAGC
TGGTGCTGGAACCTGCGCAGAGGCGGCCCTGGAGGGGCTACGCTACACGGCAGTGCTG
AAGCAGCAGCAACGCAGCACTCCATGGCCCTGCACTGGGGGGCGCTGTGGCCCAGCC
CGCCAGCCCATGTGGGGCCTGAGGGACACTCCCATCCCCGCTGGAAACTGTCCA
GCGCCGAGACATATTCACGCATGCGTCTGAAGCTGGTGCCCAACCATCACTTCGACCCTCAC
CTGGAAGCCAGCGCTCTCCGAGACAATCTGGGTGAGGTTCCCCTGACACCACCACCAGCAGGAGGC CTCACTGCCTCTGGCAGTGACCAAAGAGGCCAAAGTGAGCACCCCACCCGAGTTGCTGCAGG AGGACCAGCTCGGCGAGGACGAGCTGGCTGAGCTGGAGACCCCGATGGAGGCAGCAGAACTG CAACTACTTCCTCAACTTCCCATGCAAGGTGGGCACGACCCCAGTCTCATCTCCTAGCCAGA CTCCGAGACCCCAGCCTGGCCCCATCCCACCCCATACCCAGGTACGGAACCAGGTGTACTCG TGGCTCCTGCGCCTACGGCCCCCTCTCAAGGCTACCTAAGCAGCCGCTCCCCCAGGAGAT GCTGCGTGCCTCAGGCCTTACCCAGAAATGGGTACAGCGTGAGATATCCAACTTCGAGTACT TGATGCAACTCAACACCATTGCGGGGCCGACCTACAATGACCTGTCTCAGTACCCTGTGTTC CCCTGGGTCCTGCAGGACTACGTGTCCCCAACCCTGGACCTCAGCAACCCAGCCGTCTTCCG GGACCTGTCTAAGCCCATCGGTGTGGTGAACCCCAAGCATGCCCAGCTCGTGAGGGAGAAGT AGCACGCCTGGAGACCGCCGATGTGACCGGCAGTTCCACTCGGTGGCGCAGCCTGGC
AGGCACGCCTGGAGACCCTGCCGATGTGAAGGAGCTCATCCCGGAATTCTTCTACTTTCCT
GACTTCCTGGAGAACCAGAACGGTTTTGACCTGGGCTGTCTCCAGCTGACCAACGAGAAGGT
AGGCGATGTGGTGCTACCCCCGTGGGCCAGCTCTCCTGAGGACTTCATCCAGCAGCACCGCC
AGGCTCTGGAGTCGGAGTATGTGTCTCCACACCTACACGAGTTGGATCGACCTCATCTTTTGGC
TACAAGCAGCGGGGGCCACCAGCAGGAGGCCCCCCAATGTCTTCTTATTACTGCACCCTATGA GGGGGCTGTAGACCTGGACCATGTGACAGATGAGCGGGAACGGAAGGCTCTGGAGGGCATTA GGGGGCTGTAGACCTGGACCATGTGACAGATGAGCGGGAACGGAAGGCTCTGGAGGGCATTA
TCAGCAACTTTGGGCAGACTCCCTGTCAGCTGCTGAAGGAGCCACATCCAACTCGGCTCTCA
GCTGAGGAAGCAGCCCATCGCCTTGCACGCCTGGACACTAACTCACCTAGCATCTTCCAGCA
CCTGGACGAACTCAAGGCATTCTTCGCAGAGGTGACTGTGAGTGCCAGTGGGCTGCTGGGCA
CCACAGCTGGTTGCCCTATGACCGCAACATAAGCAACTACTTCAGCTTCAGCAAAGACCCC
ACCATGGGCAGCCACAAGACGCAGCGACTGCTGAGTGGCCCGTGGGTGCCAGGCAGTGGTG
GAGTGGACAAGCACTGCAGTGGCCCCGGATGGAAAGCTGCTATTCAGCGGTGGCCACTGGGCATGTGATGTAGCCAGCTCACCTCACCTCACCTCAGCTCCCG
GGACACCACGTGCATGGTGGCGCTCCTGCATCAGGTGTCTACCTCAGTAGGCCTGCCAC
CCAAAGCCTGTGCAGGTCCTGTATGGGCATCTGAGCTGTGAGCCTTTGGCCATCAGCACC
GAACCTTGTGCAGGTCCTTTGCATTCTGAGCATCTGATCATACACACCTGTACCCCG
GAACCTTTGACATGGCTTCTTGGATCTGAGCACCCTGTGATCATCACACACTGTACCCCG GAACTTGACATGGCTGTGTCTGGATCTGAGGATGGAACTGTGATCATACACACTGTACGCCG CGGACAGTTTGTAGCGGCACTACGGCCTCTGGGTGCCACATTCCCTGGACCTATTTTCCACC TGGCATTGGGGTCCGAAGGCCAGATTGTGGTACAGAGCTCAGCGTGGGAACGTCCTGGGGCC CAGGTCACCTACTCCTTGCACCTGTATTCAGTCAATGGGAAGTTGCGGGCTTCACTGCCCCT GGCAGAGCAGCCTACAGCCCTGACGGTGACAGAGGACTTTGTGTTGCTGGGCACCGCCCAGT CCTACTGAGGCGCGCTGAACCTGGCCAGTCCGGCTGCTCGGGCCCCCCCGGCAGGCCTGGCCCGGGAGGCCCCGCCCAGAAGTCGGCGGGAACACCCCGGGGTGGGCAGCCCAGGGGGTGA GCGGGGCCCACCCTGCCCAGCTCAGGGATTGGCGGCGATGTTACCCCCTCAGGGATTGGCG GGCGGAAGTCCCGCCCTCGCCGGCTGAGGGGCCGCCCTGAGGGCCAGCACTGGCGTCT

FIGURE 23

MSQFEMDTYAKSHDLMSGFWNACYDMLMSSGQRRQWERAQSRRAFQELVLEPAQRRARLEGL RYTAVLKOOATOHSMALLHWGALWRQLASPCGAWALRDTPIPRWKLSSAETYSRMRLKLVPN HHFDPHLEASALRDNLGEVPLTPTEEASLPLAVTKEAKVSTPPELLQEDQLGEDELAELETP MEAAELDEQREKLVLSAECQLVTVVAVVPGLLEVTTQNVYFYDGSTERVETEEGIGYDFRRP LAOLREVHLRRFNLRRSALELFFIDOANYFLNFPCKVGTTPVSSPSQTPRPQPGPIPPHTQV RNOVYSWLLRLRPPSQGYLSSRSPQEMLRASGLTQKWVQREISNFEYLMQLNTIAGRTYNDL SQYPVFPWVLQDYVSPTLDLSNPAVFRDLSKPIGVVNPKHAQLVREKYESFEDPAGTIDKFH YGTHYSNAAGVMHYLIRVEPFTSLHVQLQSGRFDCSDRQFHSVAAAWQARLESPADVKELIP EFFYFPDFLENONGFDLGCLOLTNEKVGDVVLPPWASSPEDFIQQHRQALESEYVSAHLHEW IDLIFGYKQRGPAAEEALNVFYYCTYEGAVDLDHVTDERERKALEGIISNFGQTPCQLLKEP HPTRLSAEEAAHRLARLDTNSPSIFOHLDELKAFFAEVTVSASGLLGTHSWLPYDRNISNYF SFSKDPTMGSHKTORLLSGPWVPGSGVSGOALAVAPDGKLLFSGGHWDGSLRVTALPRGKLL SOLSCHLDVVTCLALDTCGIYLISGSRDTTCMVWRLLHQGGLSVGLAPKPVQVLYGHGAAVS CVAISTELDMAVSGSEDGTVIIHTVRRGQFVAALRPLGATFPGPIFHLALGSEGQIVVQSSA WERPGAOVTYSLHLYSVNGKLRASLPLAEQPTALTVTEDFVLLGTAQCALHILQLNTLLPAA PPLPMKVAIRSVAVTKERSHVLVGLEDGKLIVVVAGQPSEVRSSQFARKLWRSSRRISQVSS GETEYNPTEAR

N-glycosylation site.

amino acids 677-681

amino acids 985-989

cAMP- and cGMP-dependent protein kinase phosphorylation site.

Tyrosine kinase phosphorylation site.

amino acids 56-65, 367-376, 543-551

N-myristoylation site.

amino acids 61-67, 436-442, 604-610, 610-616, 664-670, 691-697, 706-712, 711-717, 769-775, 785-791, 802-808, 820-826, 834-840, 873-879, 912-918, 954-960

FIGURE 24

CACGGCCCACCTTGTGAACTCCTCGTGCCCAGGGCTGATGTGCGTCTTCCAGGGCTACTCAT CCAAAGGCCTAATCCAACGTTCTGTCTTCAATCTGCAAATCTATGGGGTCCTGGGGCTCTTC TGGACCCTTAACTGGGTACTGGCCCTGGGCCAATGCGTCCTCGCTGGAGCCTTTGCCTCCTT CTACTGGGCCTTCCACAAGCCCCAGGACATCCCTACCTTCCCCTTAATCTCTGCCTTCATCC GCACACTCCGTTACCACACTGGGTCATTGGCATTTGGAGCCCTCATCCTGACCCTTGTGCAG ATAGCCCGGGTCATCTTGGAGTATATTGACCACAAGCTCAGAGGAGTGCAGAACCCTGTAGC CCGCTGCATCATGTGCTGTTTCAAGTGCTGCCTCTGGTGTCTGGAAAAATTTATCAAGTTCC TAAACCGCAATGCATACATCATGATCGCCATCTACGGGAAGAATTTCTGTGTCTCAGCCAAA AATGCGTTCATGCTACTCATGCGAAACATTGTCAGGGTGGTCGTCCTGGACAAAGTCACAGA CCTGCTGCTGTTCTTTGGGAAGCTGCTGGTGGTCGGAGGCGTGGGGGGTCCTGTCCTTCTTT TTTTCTCCGGTCGCATCCCGGGGCTGGGTAAAGACTTTAAGAGCCCCCCACCTCAACTATTAC TGGCTGCCCATCATGACCTCCATCCTGGGGGGCCTATGTCATCGCCAGCGGCTTCTTCAGCGT TTTCGGCATGTGTGGGACACGCTCTTCCTCTGCTTCCTGGAAGACCTGGAGCGGAACAACG GCTCCCTGGACCGGCCCTACTACATGTCCAAGAGCCTTCTAAAGATTCTGGGCAAGAAGAAC GAGGCGCCCCGGACAACAAGAAGAGGAAGAAGTGACAGCTCCGGCCCTGATCCAGGACTGC ACCCCACCCCACCGTCCAGCCATCCAACCTCACTTCGCCTTACAGGTCTCCATTTTGTGGT AAAAAAGGTTTTAGGCCAGGCGCCGTGGCTCACGCCTGTAATCCAACACTTTGAGAGGCTG AGGCGGCGGATCACCTGAGTCAGGAGTTCGAGACCAGCCTGGCCAACATGGTGAAACCTCC GTCTCTATTAAAAATACAAAAATTAGCCGAGAGTGGTGGCATGCACCTGTCATCCCAGCTAC TCGGGAGGCTGAGGCAGGAGATCGCTTGAACCCGGGAGGCAGAGGTTGCAGTGAGCCGAGA AAAGATTTTATTAAAGATATTTTGTTAACTC

FIGURE 25

RTRGRTRGGCEKVPINTSCNPTAHLVNSSCPGLMCVFQGYSSKGLIQRSVFNLQIYGVLGLF
WTLNWVLALGQCVLAGAFASFYWAFHKPQDIPTFPLISAFIRTLRYHTGSLAFGALILTLVQ
IARVILEYIDHKLRGVQNPVARCIMCCFKCCLWCLEKFIKFLNRNAYIMIAIYGKNFCVSAK
NAFMLLMRNIVRVVVLDKVTDLLLFFGKLLVVGGVGVLSFFFFSGRIPGLGKDFKSPHLNYY
WLPIMTSILGAYVIASGFFSVFGMCVDTLFLCFLEDLERNNGSLDRPYYMSKSLLKILGKKN
EAPPDNKKRKK

FIGURE 26

GAGTCTTGACCGCCGGGCTCTTGGTACCTCAGCGCGAGCGCCAGGCGTCCGGCCGCCGT GGCTATGTTCGTGTCCGATTTCCGCAAAGAGTTCTACGAGGTGGTCCAGAGCCAGAGGGTCC CAGTGTGACCACGTGCAATATACGCTGGTTCCAGTTTCTGGGTGGCAAGAACTTGAAACTGC ATTTCTTGAGCATAAAGAACAGTTTCATTATTTTATTCTCATAAACTGTGGAGCTAATGTAG CCAGTCAATGTCGTCAATGTATACAACGATACCCAGATCAAATTACTCATTAAACAAGATGA TGACCTTGAAGTTCCCGCCTATGAAGACATCTTCAGGGATGAAGAGGAGGATGAAGAGCATT CAGGAAATGACAGTGATGGGTCAGAGCCTTCTGAGAAGCGCACACGGTTAGAAGAGGAGATA GTGGAGCAAACCATGCGGAGGGGCGGCGAGAGTGGGAGGCCCGGAGAAGAGACATCCT $\verb|CTTTGACTACGAGCAGTATGAATATCATGGGACATCGTCAGCCATGGTGATGTTTGAGCTGG|\\$ CTTGGATGCTGTCCAAGGACCTGAATGACATGCTGTGGTGGGCCATCGTTGGACTAACAGAC CAGTGGGTGCAAGACAAGATCACTCAAATGAAATACGTGACTGATGTTGGTGTCCTGCAGCG CCACGTTTCCCGCCACAACCACCGGAACGAGGATGAGGAGAACACACTCTCCGTGGACTGCA CACGGATCTCCTTTGAGTATGACCTCCGCCTGGTGCTCTACCAGCACTGGTCCCTCCATGAC AGCCTGTGCAACACCAGCTATACCGCAGCCAGGTTCAAGCTGTGGTCTGTGCATGGACAGAA GCGGCTCCAGGAGTTCCTTGCAGACATGGGTCTTCCCCTGAAGCAGGTGAAGCAGAAGTTCC AGGCCATGGACATCTCCTTGAAGGAGAATTTGCGGGAAATGATTGAAGAGTCTGCAAATAAA TCTGGCCAGCGACGTGTCTTTGCCACCATGTCTTTGATGGAGAGCCCCGAGAAGGATGGCT CAGGGACAGATCACTTCATCCAGGCTCTGGACAGCCTCTCCAGGAGTAACCTGGACAAGCTG TACCATGGCCTGGAACTCGCCAAGAAGCAGCTGCGAGCCACCCAGCAGACCATTGCCAGCTGC CTTTGCACCAACCTCGTCATCTCCCAGGGGCCTTTCCTGTACTGCTCTCTCATGGAGGGCAC TCCAGATGTCATGCTGTTCTCTAGGCCGGCATCCCTAAGCCTGCTCAGCAAACACCTGCTCA AGTCCTTTGTGTGTTCGACAAAGAACCGGCGCTGCAAACTGCTGCCCCTGGTGATGGCTGCC CCCCTGAGCATGGCACAGTGACCGTGGTGGGCATCCCCCCAGAGACCGACAGCTC GGACAGGAAGACTTTTTTGGGAGGGCGTTTGAGAAGGCAGCGGAAAGCACCAGCTCCCGGA TGCTGCACAACCATTTTGACCTCTCAGTAATTGAGCTGAAAGCTGAGGATCGGAGCAAGTTT $\tt CTGGACGCACTTATTTCCCTCCTGTCC\textbf{TAG}GAATTTGATTCTTCCAGAATGACCTTCTTATT$ TATGTAACTGGCTTTCATTTAGATTGTAAGTTATGGACATGATTTGAGATGTAGAAGCCATT TTTTATTAAATAAAATGCTTATTTTAGGAAA

FIGURE 27

MFVSDFRKEFYEVVQSQRVLLFVASDVDALCACKILQALFQCDHVQYTLVPVSGWQELETAF
LEHKEQFHYFILINCGANVDLLDILQPDEDTIFFVCDSHRPVNVVNVYNDTQIKLLIKQDDD
LEVPAYEDIFRDEEEDEEHSGNDSDGSEPSEKRTRLEEEIVEQTMRRRQRREWEARRRDILF
DYEQYEYHGTSSAMVMFELAWMLSKDLNDMLWWAIVGLTDQWVQDKITQMKYVTDVGVLQRH
VSRHNHRNEDEENTLSVDCTRISFEYDLRLVLYQHWSLHDSLCNTSYTAARFKLWSVHGQKR
LQEFLADMGLPLKQVKQKFQAMDISLKENLREMIEESANKFGMKDMRVQTFSIHFGFKHKFL
ASDVVFATMSLMESPEKDGSGTDHFIQALDSLSRSNLDKLYHGLELAKKQLRATQQTIASCL
CTNLVISQGPFLYCSLMEGTPDVMLFSRPASLSLLSKHLLKSFVCSTKNRRCKLLPLVMAAP
LSMEHGTVTVVGIPPETDSSDRKNFFGRAFEKAAESTSSRMLHNHFDLSVIELKAEDRSKFL
DALISLLS

FIGURE 28

FIGURE 29

CAGGAACCCTCTCTTTGGGTCTGGATTGGGACCCCTTTCCAGTACCATTTTTTCTAGTGAAC CACGAAGGGACGATACCAGAAAACACCCTCAACCCAAAGGAAATAGACTACAGCCCCAATTG GTCTTCCCTTTATCGAGTCAAGAAACCCCCCCTTCTTGAGCTATTTACAGCTTTTAACAATT GAGTAAAGTACGCTCCGGTCACC**ATG**GTGACAGCCGCCCTGGGTCCCGTCTGGGCAGCGCTC CTGCTCTTTCTCCTGATGTGTGAGATCCGTATGGTGGAGCTCACCTTTGACAGAGCTGTGGC CTGCTCTTTCTCCTGATGTGTGAGATCCGTATGGTGGAGCTCACCTTTGACAGAGCTGTGGC
CAGCGGCTGCCAACGGTGCTGTGACTCTGAGGACCCCTGGATCCTGCCCATGTATCCTCAG
CCTCTTCCTCCGGCCGCCCCACGCCCTGCCTGAGATCAGACCCTACATTAATATCACCATC
CTGAAGGGTGACAAAGGGGACCCAGGCCCAATGGGCCTGCCAGGGTACATGGGCAGGAGGG
TCCCCAAGGGGAGCCTTGCCCAAGGGCAAAGGGTGACAAGGGGAGATGGGCAGCCCCG
GCGCCCCGTGCCAGAAGCGCTTCTTCGCCTTCTCAGTGGGCCCAAACGGCCCTGCACAGC
GCGCAAGAGCGCTGCTCTCCAAAGGGTCTTTTGTAACCTTTGATGGGTGCTTTTGA CATGGCGACCGGCCAGTTTGCTGCTCCCCTGCGTGGCATCTACTTCTTCAGCCTCAATGTGC ACAGCTGGAATTACAAGGAGACGTACGTGCACATTATGCATAACCAGAAAGAGGCTGTCATC CTGTACGCGCAGCCAGCGAGCGCAGCATCATGCAGAGCCAGAGTGTGATGCTGGACCTGGCCTACGGGGACCGCGTCTGGGTGCGGCTCTTCAAGCGCCAGCGCGAGAACGCCATCTACAGCA ATGGGGTCAGTCTATGAACCACACTGGAGCAACCAGGTTCTAGGACTTTCTCAATATTCTAG GTCCAGCCCAGGCGGGAGAGATGTGTACATAGGTTTTAAAGCAGACCCAGAGCTCATGGGG AGCTAGAGCAAAGGAGGGACCTCAGGCCTTCCGTTTCTTCTTCCAGGGTGGGGTGGCCTGGT

FIGURE 30

MVTAALGPVWAALLLFLLMCEIRMVELTFDRAVASGCQRCCDSEDPLDPAHVSSASSSGRPH
ALPEIRPYINITILKGDKGDPGPMGLPGYMGREGPQGEPGPQGSKGDKGEMGSPGAPCQKRF
FAFSVGRKTALHSGEDFQTLLFERVFVNLDGCFDMATGQFAAPLRGIYFFSLNVHSWNYKET
YVHIMHNQKEAVILYAQPSERSIMQSQSVMLDLAYGDRVWVRLFKRQRENAIYSNDFDTYIT
FSGHLIKAEDD

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 72-75

Clq domain proteins.

amino acids 144-178, 78-111 and 84-117

FIGURE 31

ACTCGAACGCAGTTGCTTCGGGACCCAGGACCCCCTCGGGCCCGACCCGCCAGGAAAGACTG AGGCCGCGGCCTGCCCGGCCGGCTCCCTGCGCCGCCGCCCTCCCGGGACAGAAGATGTG CTCCAGGGTCCCTCTGCTGCTGCTGCTGCTGCTACTGGCCCTGGGGCCTGGGGTGCAGG GCTGCCCATCCGGCTGCCAGTGCAGCCAGCCAGAGACTCTTCTGCACTGCCCGCCAGGGG ACCACGGTGCCCGAGACGTGCCACCCGACACGGTGGGGCTGTACGTCTTTGAGAACGGCAT CACCATGCTCGACGCAGGCAGCTTTGCCGGCCTGCCGGGCCTGCAGCTCCTGGACCTGTCAC AGAACCAGATCGCCAGCCTGCCCAGCGGGGTCTTCCAGCCACTCGCCAACCTCAGCAACCTG GACCTGACGGCCAACAGGCTGCATGAAATCACCAATGAGACCTTCCGTGGCCTGCGGCGCCCT CGAGCGCCTCTACCTGGGCAAGAACCGCATCCGCCACATCCAGCCTGGTGCCTTCGACACGC TCGACCGCCTCCTGGAGCTCAAGCTGCAGGACAACGAGCTGCGGGCACTGCCCCCGCTGCGC CTGCCCGCCTGCTGCTGGACCTCAGCCACAACAGCCTCCTGGCCCTGGAGCCCGGCAT CCTGGACACTGCCAACGTGGAGGCGCTGCGGCTGGCTGGTCTGGGGCTGCAGCAGCTGGACG AGGGGCTCTTCAGCCGCTTGCGCAACCTCCACGACCTGGATGTGTCCGACAACCAGCTGGAG CGAGTGCCACCTGTGATCCGAGGCCTCCGGGGCCTGACGCCTGCGGCTGGCCGGCAACAC CCGCATTGCCCAGCTGCGGCCCGAGGACCTGGCCGGCCTGGCTGCCCTGCAGGAGCTGGATG TGAGCAACCTAAGCCTGCAGGCCCTGCCTGGCGACCTCTCGGGCCTCTTCCCCCGCCTGCGG CTGCTGGCAGCTGCCCGCAACCCCTTCAACTGCGTGTGCCCCCTGAGCTGGTTTGGCCCCTG GGTGCGCGAGAGCCACGTCACACTGGCCAGCCCTGAGGAGACGCGCTGCCACTTCCCGCCCA AGAACGCTGGCCGGCTGCTCCTGGAGCTTGACTACGCCGACTTTGGCTGCCCAGCCACCACC ACCACAGCCACAGTGCCCACCACGAGGCCCGTGGTGCGGGAGCCCACAGCCTTGTCTTCTAG CTGCCCCACCGACTGTAGGGCCTGTCCCCCAGCCCCAGGACTGCCCACCGTCCACCTGCCTC AATGGGGGCACATGCCACCTGGGGACACGGCACCACCTGGCGTGCTTGTGCCCCGAAGGCTT CACGGGCCTGTACTGTGAGAGCCAGATGGGGCAGGGGACACGGCCCAGCCCTACACCAGTCA CGCCGAGGCCACCACGGTCCCTGACCCTGGGCATCGAGCCGGTGAGCCCCACCTCCCTGCGC GTGGGGCTGCAGCGCTACCTCCAGGGGAGCTCCGTGCAGCTCAGGAGCCTCCGTCTCACCTA AGTACACGGTCACCCAGCTGCGGCCCAACGCCACTTACTCCGTCTGTGTCATGCCTTTGGGG CCCGGGCGGTGCCGAGGGCGAGGAGGCCTGCGGGGAGGCCCATACACCCCCAGCCGTCCA CTCCAACCACGCCCCAGTCACCCAGGCCCGCGAGGGCAACCTGCCGCTCCTCATTGCGCCCG CCCTGGCCGCGGTGCTCCTGGCCGCGCTGGCTGCGGTGGGGGCAGCCTACTGTGTGCGGCGG GGGCGGCCATGGCAGCAGCGGCTCAGGACAAAGGGCAGGTGGGGCCAGGGGCTGGGCCCCT GGAACTGGAGGGAGTGAAGGTCCCCTTGGAGCCAGGCCCGAAGGCAACAGAGGGCGGTGGAG AGGCCCTGCCCAGCGGGTCTGAGTGTGAGGTGCCACTCATGGGCTTCCCAGGGCCTGGCCTC CAGTCACCCCTCCACGCAAAGCCCTACATC**TAA**GCCAGAGAGAGACAGGGCAGCTGGGGCCG GGCTCTCAGCCAGTGAGATGGCCAGCCCCCTCCTGCTGCCACACCACGTAAGTTCTCAGTCC CAACCTCGGGGATGTGTGCAGACAGGGCTGTGTGACCACAGCTGGGCCCTGTTCCCTCTGGA CCTCGGTCTCCTCATCTGTGAGATGCTGTGGCCCAGCTGACGAGCCCTAACGTCCCCAGAAC CGAGTGCCTATGAGGACAGTGTCCGCCTGCCCTCCGCAACGTGCAGTCCCTGGGCACGGCG GGCCCTGCCATGTGCTGGTAACGCATGCCTGGGTCCTGCTGGGCTCTCCCACTCCAGGCGGA CCCTGGGGGCCAGTGAAGGAAGCTCCCGGAAAGAGCAGAGGGAGAGCGGGTAGGCGGCTGTG TGACTCTAGTCTTGGCCCCAGGAAGCGAAGGAACAAAAGAAACTGGAAAGGAAGATGCTTTA GGAACATGTTTTGCTTTTTTAAAATATATATATTTTATAAGAGATCCTTTCCCATTTATTCTG GGAAGATGTTTTTCAAACTCAGAGACAAGGACTTTGGTTTTTGTAAGACAAACGATGATATG AAGGCCTTTTGTAAGAAAAAATAAAAGATGAAGTGTGAAA

FIGURE 32

MCSRVPLLLPLLLLALGPGVQGCPSGCQCSQPQTVFCTARQGTTVPRDVPPDTVGLYVFEN GITMLDAGSFAGLPGLQLLDLSQNQIASLPSGVFQPLANLSNLDLTANRLHEITNETFRGLR RLERLYLGKNRIRHIQPGAFDTLDRLLELKLQDNELRALPPLRLPRLLLLDLSHNSLLALEP GILDTANVEALRLAGLGLQQLDEGLFSRLRNLHDLDVSDNQLERVPPVIRGLRGLTRLRLAG NTRIAQLRPEDLAGLAALQELDVSNLSLQALPGDLSGLFPRLRLLAAARNPFNCVCPLSWFG PWVRESHVTLASPEETRCHFPPKNAGRLLLELDYADFGCPATTTTATVPTTRPVVREPTALS SSLAPTWLSPTAPATEAPSPPSTAPPTVGPVPQPQDCPPSTCLNGGTCHLGTRHHLACLCPE GFTGLYCESQMGQGTRPSPTPVTPRPPRSLTLGIEPVSPTSLRVGLQRYLQGSSVQLRSLRL TYRNLSGPDKRLVTLRLPASLAEYTVTQLRPNATYSVCVMPLGPGRVPEGEEACGEAHTPPA VHSNHAPVTQAREGNLPLLIAPALAAVLLAALAAVGAAYCVRRGRAMAAAAQDKGQVGPGAG PLELEGVKVPLEPGPKATEGGGEALPSGSECEVPLMGFPGPGLOSPLHAKPYI

FIGURE 33

GAATCATCCACGCACCTGCAGCTCTGCTGAGAGAGTGCAAGCCGTGGGGGTTTTGAGCTCAT $\verb|CTTCATCATTCATATGAGGAAATAAGTGGTAAAATCCTTGGAAATACAATG| AGACTCATCAG|$ CTTCATCATTCATATGAGGAAATAAGTGGTAAAATCCTTGGAAATACAATGAGAGCTCATCAG
AAACATTTACATATTTTGTAGTATTGTTATGACAGCAGAGGGTGATGCTCCAGAGCTGCCAG
AAGAAAGGGAACTGATGACCAACTGCTCCAACATGTCTCTAAGAAAGGTTCCCGCAGACTTG
ACCCCAGCCACAACGACACTGGATTTATCCTATAACCTCCTTTTTCAACTCCAGAGTTCAGA
TTTTCATTCTGTCTCCAAACTGAGAGTTTTGATTCTATGCCATAACAGAATTCAACAGCTGG
ATCTCAAAACCTTTGAATTCAACAAGGAGTTAAGATATTTAGATTTTGTCTAATAACAGACTG
AAGAGTGTAACTTGGTATTTACTGGCAGGTCTCAGGTATTTAGATCTTTTTTAATGACTT
TGACACCATGCCTATCTGTGAGGAAGCTTGCACATGTCACCCTCATTTAAACACACA TTCTTAGGATTCAGAACTCTTCCTCATTATGAAGAAGGTAGCCTGCCCATCTTAAACACAAC AAAACTGCACATTGTTTTACCAATGGACACAAATTTCTGGGTTCTTTTTGCGTGATGGAATCA AGACTTCAAAAATATTAGAAATGACAAATATAGATGGCCAAAAGCCCAATTTGTAAGTTATGAA ATGCAACGAAATCTTAGTTTAGAAAATGCTAAGACATCGGTTCTATTGCTTAATAAAGTTGA CACAAATGCCACACATGCTTTTCCCGAATTATCCTACGAAATTCCAATATTTAAATTTTGCC AATAATATCTTAACAGACGAGTTGTTTAAAAGAACTATCCAACTGCCTCACTTGAAAACTCT CATTTTGAATGGCAATAAACTGGAGACACTTTCTTTAGTAAGTTGCTTTGCTAACAACACAC CCTTGGAACACTTGGATCTGAGTCAAAATCTATTACAACATAAAAATGATGAAAATTGCTCA TGGCCAGAAACTGTGGTCAATATGAATCTGTCATACAATAAATTGTCTGATTCTGTCTTCAG GTGCTTGCCCAAAAGTATTCAAATACTTGACCTAAATAATAACCAAATCCAAACTGTACCTA AAGAGACTATTCATCTGATGGCCTTACGAGAACTAAATATTGCATTTAATTTTCTAACTGAT CTCCCTGGATGCAGTCATTTCAGTAGACTTTCAGTTCTGAACATTGAAATGAACTTCATTCTCAGCCCATCTCTGGATTTTGTTCAGAGCTGCCAGGAAGTTAAAACTCTAAATGCGGGAAGAA GTTAAAAGACGTTCATCTCCACGAATTATCTTGCAACACACGCTCTGTTGATTGTCACCATTG TGGTTATTATGCTAGTTCTGGGGTTGGCTGTGGCCTTCTGCTGTCTCCACTTTGATCTGCCC TGGTATCTCAGGATGCTAGGTCAATGCACACAAACATGGCACAGGGTTAGGAAAACAACCCA ACCTTCGAGCTGCTATTAATGTTAATGTATTAGCCACCAGAGAAATGTATGAACTGCAGACA TTCACAGAGTTAAATGAAGAGTCTCGAGGTTCTACAATCTCTCTGATGAGAACAGATTGTCT A**TAA**AATCCCACAGTCCTTGGGAAGTTGGGGACCACATACACTGTTGGGATGTACATTGATA CAACCTTTATGATGGCAATTTGACAATATTTATTAAAATAAAAAATGGTTATTCCCTTCATA TCAGTTTCTAGAAGGATTTCTAAGAATGTATCCTATAGAAACACCTTCACAAGTTTATAAGG GCTTATGGAAAAAGGTGTTCATCCCAGGATTGTTTATAATCATGAAAAATGTGGCCAGGTGC GAGATGGAGACCATCCTGGCCAACATGGTGAAACCCTGTCTCTACTAAAAATACAAAAATTA GCTGGGCGTGATGGTGCACGCCTGTAGTCCCAGCTACTTGGGAGGCTGAGGCAGGAGAATCG CTTGAACCCGGGAGGTGGCAGTTGCAGTGAGCTGAGATCGAGCCACTGCACTCCAGCCTGGT TCATGGCCACAAAATAAGGTCTAATTCAATAAATTATAGTACATTAATGTAATATAATATATTA CATGCCACTAAAAGAATAAGGTAGCTGTATATTTCCTGGTATGGAAAAAACATATTAATAT GTTATAAACTATTAGGTTGGTGCAAAACTAATTGTGGTTTTTTGCCATTGAAATGGCATTGAA ATAAAAGTGTAAAGAAATCTATACCAGATGTAGTAACAGTGGTTTGGGTCTGGGAGGTTGGA TTACAGGGAGCATTTGATTTCTATGTTGTGTATTTCTATAATGTTTGAATTGTTTAGAATGA ATCTGTATTTCTTTTATAAGTAGAAAAAAAAAAAAAAGATAGTTTTTACAGCCT

FIGURE 34

MRLIRNIYIFCSIVMTAEGDAPELPEERELMTNCSNMSLRKVPADLTPATTTLDLSYNLLFQ
LQSSDFHSVSKLRVLILCHNRIQQLDLKTFEFNKELRYLDLSNNRLKSVTWYLLAGLRYLDL
SFNDFDTMPICEEAGNMSHLEILGLSGAKIQKSDFQKIAHLHLNTVFLGFRTLPHYEEGSLP
ILNTTKLHIVLPMDTNFWVLLRDGIKTSKILEMTNIDGKSQFVSYEMQRNLSLENAKTSVLL
LNKVDLLWDDLFLILQFVWHTSVEHFQIRNVTFGGKAYLDHNSFDYSNTVMRTIKLEHVHFR
VFYIQQDKIYLLLTKMDIENLTISNAQMPHMLFPNYPTKFQYLNFANNILTDELFKRTIQLP
HLKTLILNGNKLETLSLVSCFANNTPLEHLDLSQNLLQHKNDENCSWPETVVNMNLSYNKLS
DSVFRCLPKSIQILDLNNNQIQTVPKETIHLMALRELNIAFNFLTDLPGCSHFSRLSVLNIE
MNFILSPSLDFVQSCQEVKTLNAGRNPFRCTCELKNFIQLETYSEVMMVGWSDSYTCEYPLN
LRGTRLKDVHLHELSCNTALLIVTIVVIMLVLGLAVAFCCLHFDLPWYLRMLGQCTQTWHRV
RKTTQEQLKRNVRFHAFISYSEHDSLWVKNELIPNLEKEDGSILICLYESYFDPGKSISENI
VSFIEKSYKSIFVLSPNFVQNEWCHYEFYFAHHNLFHENSDHIILILLEPIPFYCIPTRYHK
LKALLEKKAYLEWPKDRRKCGLFWANLRAAINVNVLATREMYELQTFTELNEESRGSTISLM
RTDCL

FIGURE 35

GGGGGCTTTCTTGGGCTTGCTTGGAACACCTGCCTCCAAGGACCGGCCTCGGAGGGGTCGCCGGGAAAGG GAGGGAAGAAGGAAGGCCGGGCCCCCCTGCGCCCCCGCGCCTCTGCGCGCCCTGTCCGCCCCGGC CTGGCAGTGACCCTGGCCGGGGTCGGAGCCCAGGGCCCTCGAGGACCCTGATTATTACGGGCAGGAGAT $\tt CCGGGGAGGGGGGGGGGGCCCGCGGGGGCCCCAGGGGGGCCCAAGAGGCCCAAGAAAGCTCCC$ AAGAGGGAGAAGTCGGCTCCGGAGCCGCCTCCACCAGGTAAACACAACAAAAAAGTTATGAGAACCAAGAG TTGGTCTGGAAACCTTAAAAATCACAGACTTCCAGCTCCATGCCTCCACGGTGAAGCGCTATGGCCTGGGGGCA AAATGACCTCCAGCAGTGGATTGAAGTGGATGCTCGGCGCCTGACCAGATTCACTGGTGTCATCACTCAAGGGA GGAACTCCCTCTGGCTGAGTGACTGGGTGACATCCTATAAGGTCATGGTGAGCAATGACAGCCACACGTGGGTC ACTGTTAAGAATGGATCTGGAGACATGATATTTGAGGGAAACAGTGAGAAGGAGATCCCTGTTCTCAATGAGCT ACCCGTCCCCATGGTGGCCCGCTACATCCGCATAAACCCTCAGTCCTGGTTTGATAATGGGAGCATCTGCATGA GAATGGAGATCCTGGGCTGCCCACTGCCAGATCCTAATAATTATTATCACCGCCGGAACGAGATGACCACCACT CAATATCACCAGAATTTACAACATTGGAAAAAGCCACCAGGGCCTGAAGCTGTATGCTGTGGAGATCTCAGATC ACCCTGGGGAGCATGAAGTCGGTGAGCCCGAGTTCCACTACATCGCGGGGGCCCCACGGCAATGAGGTGCTGGGC $\tt CGGGAGCTGCTGCTGCTGCTGCTGCTGTTCAGGAGTACTTGGCCCGGAATGCGCGCATCGTCCACCT$ GGTGGAGGAGACGCGGATTCACGTCCTCCCTCAACCCCGATGGCTACGAGAAGGCCTACGAAGGGGGCT $\tt CGGAGCTGGGAGGCTGGTCCCTGGGACGCTGGACCCACGATGGAATTGACATCAACAACAACTTTCCTGATTTA$ AACACGCTGCTCTGGGAGGCAGAGGATCGACAGAATGTCCCCAGGAAAGTTCCCAATCACTATATTGCAATCCC ACACCGCCTCATGACAGACGCCCGGAGGAGGGTGTGCCACACGGAGGACTTCCAGAAGGAGGAGGACTGTCA ATGGGGCCTCCTGGCACACCGTCGCTGGAAGTCTGAACGATTTCAGCTACCTTCATACAAACTGCTTCGAACTG TCCATCTACGTGGGCTGTGATAAATACCCACATGAGAGCCAGCTGCCCGAGGAGTGGGAGAATAACCGGGAATC TCTGATCGTGTTCATGGAGCAGGTTCATCGTGGCATTAAAGGCTTGGTGAGAGATTCACATGGAAAAGGAATCC CAAACGCCATTATCTCCGTAGAAGGCATTAACCATGACATCCGAACAGCCAACGATGGGGATTACTGGCGCCTC CTGAACCCTGGAGAGTATGTGGTCACAGCAAAGGCCGAAGGTTTCACTGCATCCACCAAGAACTGTATGGTTGG TGGACTCACTCACTGTTTTCCTCTGTAATTCAAGAAGTGCCTGGAAGAGGGGTGCATTGTGAGGCAGGTCC CAAAAGGGAAGGCTGGAGGCTGTTTTCTTTTTTTTTTCCAATTATCCAAATAACTTGGACAGAGCA GAGCCTGTCCGTTCAGAGCCTCTGGCTGCATAGAAAAGGATTCTGGTGCTTCCCCTGTTTGCGTGGCAGCAAGG GTTCCACGTGCATTTGCAATTTGCACAGCTAAAATTGCAGCATTTCCCCAGCTGGGCTGTCCCAAATGTTACCA TTTGAGATGCTCCCAGGCGTCCTAAGAGAATCCACCCTCTCTGGCCCTGGGACATTGCAAGCTGCTACAAATAA ATTCTGTGTTCTTTTGACAATAGCGTCATTGCCAAGTGCACATCAGTGAGCCTCTTGAATCTGTTTAGTCTCCT TGGAGCTTCTTGCACAAATTCTGGGTCCATAAACAACCCCCAAAGTCCCTGCTGATCCAGTAGCCCTGGAGGTT $\tt CCCCAGGTAGGGAGGCCAGGCCTTCCTGAAGGGCCAGAAATTTAGCCTGGATCTCCTCTTTTAC$ GAATTGAGTGCTCATGGGTTGGCCTCATATCAGCCTGGGAGTTATTTTTGATATGTAGAATGCCAGATCTTCCA GATTAGGCTAAATGTAATGAAAACCTCTTAGGATTATCTGTGGAGCATCAGTTTGGGAAGAATTATTGAATTAT

FIGURE 36

MSRPGTATPALALVLLAVTLAGVGAQGAALEDPDYYGQEIWSREPYYARPEPELETFSPPLP
AGPGEEWERRPQEPRPPKRATKPKKAPKREKSAPEPPPPGKHSNKKVMRTKSSEKAANDDHS
VRVAREDVRESCPPLGLETLKITDFQLHASTVKRYGLGAHRGRLNIQAGINENDFYDGAWCA
GRNDLQQWIEVDARRLTRFTGVITQGRNSLWLSDWVTSYKVMVSNDSHTWVTVKNGSGDMIF
EGNSEKEIPVLNELPVPMVARYIRINPQSWFDNGSICMRMEILGCPLPDPNNYYHRRNEMTT
TDDLDFKHHNYKEMRQLMKVVNEMCPNITRIYNIGKSHQGLKLYAVEISDHPGEHEVGEPEF
HYIAGAHGNEVLGRELLLLLVQFVCQEYLARNARIVHLVEETRIHVLPSLNPDGYEKAYEGG
SELGGWSLGRWTHDGIDINNNFPDLNTLLWEAEDRQNVPRKVPNHYIAIPEWFLSENATVAA
ETRAVIAWMEKIPFVLGGNLQGGELVVAYPYDLVRSPWKTQEHTPTPDDHVFRWLAYSYAST
HRLMTDARRRVCHTEDFQKEEGTVNGASWHTVAGSLNDFSYLHTNCFELSIYVGCDKYPHES
QLPEEWENNRESLIVFMEQVHRGIKGLVRDSHGKGIPNAIISVEGINHDIRTANDGDYWRLL
NPGEYVVTAKAEGFTASTKNCMVGYDMGATRCDFTLSKTNMARIREIMEKFGKQPVSLPARR
LKLRGRKRRQRG

FIGURE 37

ATTTGGGGGATGTGGGACCTCCAATTCCCAGCCCCGGCTTCAGCTCTTTCCCAGGTGTTGACTCCAGCTCCAGC TTCAGCTCCAGCTCCAGGTCGGGCTCCAGCTCCAGCCGCAGCTTAGGCAGCGGAGGTTCTGTGTCCCAGTTGTT TTCCAATTTCACCGGCTCCGTGGATGACCGTGGGACCTGCCAGTGCTCTGTTTCCCTGCCAGACACCACCTTTC GTGAGGGAATATGTCCAATTAATTAGTGTGTATGAAAAGAAACTGTTAAACCTAACTGTCCGAATTGACATCAT GGAGAAGGATACCATTTCTTACACTGAACTGGACTTCGAGCTGATCAAGGTAGAAGTGAAGGAGATGGAAAAAAC TGGTCATACAGCTGAAGGAGAGTTTTGGTGGAAGCTCAGAAATTGTTGACCAGCTGGAGGTGGAGATAAGAAAT TCTGAAGACCAAGCTGAAAGAGTGTGAGGCCTCTAAAGATCAAAACACCCCTGTCGTCCACCCTCCCACTC ${\tt CAGGGAGCTGTGGTGTGTGTGTGAACATCAGCAAACCGTCTGTGGTTCAGCTCAACTGGAGAGGGTTT}$ ATTGAATACAGATGGGAGACTGTTGGAGTATTATAGACTGTACAACACACTGGATGATTTGCTATTGTATATAA ATGCTCGAGAGTTGCGGATCACCTATGGCCAAGGTAGTGGTACAGCAGTTTACAACAACAACATGTACGTCAAC ATGTACAACACGGGAATATTGCCAGAGTTAACCTGACCACCAACACGATTGCTGTGACTCAAACTCTCCCTAA TGCTGCCTATAATAACCGCTTTTCATATGCTAATGTTGCTTGGCAAGATATTGACTTTGCTGTGGATGAGAATG GATTGTGGGTTATTTATTCAACTGAAGCCAGCACTGGTAACATGGTGATTAGTAAACTCAATGACACCACACTT CAGGTGCTAAACACTTGGTATACCAAGCAGTATAAACCATCTGCTTCTAACGCCTTCATGGTATGTGGGGTTCT GTATGCCACCCGTACTATGAACACCAGAACAGAAGAGATTTTTTACTATTATGACACAAACACAGGGAAAGAG GCAAACTAGACATTGTAATGCATAAGATGCAGGAAAAAGTGCAGAGCATTAACTATAACCCTTTTGACCAGAAA $\texttt{CTTTATGTCTATAACGATGGTTACCTTCTGAATTATGATCTTTCTGTCTTGCAGAAGCCCCAG} \underline{\textbf{TAA}} \texttt{GCTGTTTA}$ $\tt CTAAAAGTGTGTTCATTTTGCAGCAATGTTTAGGTGCATAGTTCTACCACACTAGAGATCTAGGACATTTGTCT$ TTGTCAGAGGTCTAGGGGCACTGTGGGCCTAGTGAAGCCTACTGTGAGGAGGCTTCACTAGAAGCCTTAAATTA GGAATTAAGGAACTTAAAACTCAGTATGGCGTCTAGGGATTCTTTGTACAGGAAATATTGCCCAATGACTAGTC GGAGCTCCTCGAGGGACCAAATCTCCAACTTTTTTTTCCCCTCACTAGCACCTGGAATGATGCTTTGTATGTGG CAGATAAGTAAATTTGGCATGCTTATATATTCTACATCTGTAAAGTGCTGAGTTTTATGGAGAGAGGCCTTTTT ATGCATTAAATTGTACATGGCAAATAAATCCCAGAAGGATCTGTAGATGAGGCACCTGCTTTTTCTTTTCTCTC AACCAGACTTACTAACCAATTCCACCCCCCCCCCCCCTTCTACTGCCTACTTTAAAAAAATTAATAGTTTT AGACTATAAGAAAATCTGATGGCAGTGACAAAGTGCTAGCATTTATTGTTATCTAATAAAGACCTTGGAGCATA TGTGCAACTTATGAGTGTATCAGTTGTTGCATGTAATTTTTTGCCTTTGTTTAAGCCTGGAACTTGTAAGAAAAT GAAAATTTAATTTTTTTTTTTTTAGGACGAGCTATAGAAAAGCTATTGAGAGTATCTAGTTAATCAGTGCAGTAGT TGGAAACCTTGCTGGTGTATGTGATGTGCTTCTGTGCTTTTGAATGACTTTATCATCTAGTCTTTGTCTATTTT

FIGURE 38

FIGURE 39

GCTCTGAAGACCAAGCTGAAAGAGTGTGAGGCCTCTAAAGATCAAACACCCCTGTCGTCCAC
CCTCCTCCCACTCCAGGGAGCTGTGGTCATGGTGGTGGTGAACATCAGCAAACCGTCTGT
GGTTCAGCTCAACTGGAGAGGGTTTTCTTATCTATATGGTGCTTGGGGTAGGGATTACTCTC
CCCAGCATCCAAACAAAGGNATGTATTGGGNGGCGCCATTGAATACAGATGGGAGACTGTTG
GAGTATTATAGACTGTACAACCCACTGGATGATTTGCTATTGTATATAAATGCTCGAGAGTT
GCGGATCACCTATGGCCAAGGTAGTGGTACAGCAGTTTACAACAACAACATGTACGTCAACA
TGTACAACACCGGGNATATTGCCAGAGTTAACCTGACC

FIGURE 40

CCGCTGCTCTTGTGACGTTGTGGAG GATCGGGGGGGCTGTGCTCCATGGCGAGCTAGAAAGGTGTTGTCCCTTGTAACATTTTGGTTGGCTATAAAGCTGTATATCGTTTGTGCTTT GGTTTGGCTATGTTCTATCTTCTCTCTCTTTACTAATGATCAAAGTGAAGAGTAGCAGTGA TCCTAGAGCTGCAGTGCACAATGGATTTTGGTTCTTTAAATTTGCTGCAGCAATTGCAATTA CGTTGGTGCTTCTGTAATGTCTATACTGCCAAAAATCCAAGAATCACAACCAAGATCTGGTT TGTTACAGTCTTCAGTAATTACAGTCTACACAATGTATTTGACATGGTCAGCTATGACCAAT GAACCAGAACCAAATTGCAACCCAAGTCTACTAAGCATAATTGGCTACAATACAACAAGCAC TGTCCCAAAGGAAGGCAGTCAGTCCAGTGGTGGCATGCTCAAGGAATTATAGGACTAATTC TCTTTTTGTTGTGTGTATTTTATTCCAGCATCCGTACTTCAAACAATAGTCAGGTTAATAAA CTGACTCTAACAAGTGATGAATCTACATTAATAGAAGATGGTGGAGCTAGAAGTGATGGATC AATCTCTTCCAGTTGGATTGGCATCGTGCTGTATGTTTGGACACTCGTGGCACCACTTGTTC TTCTCCAGTGTTCTGGCATGAATTAGATTTTACTGCTTGTCATTTTGTTATTTTCTTACCAA GTGCATTGATATGTGAAGTAGAATGAATTGCAGAGGAAAGTTTTATGAATATGGTGATGAGT GGAAATGTTTGGAACTCTGAAGGATTTAGACAAGGTTTTGAAAAGGATAATCATGGGTTAGA AGGAAGTGTTTTGAAAGTCACTTTGAAAGTTAGTTTTTGGGCCCAGCACGGTAGCTCACCCTT GGTAATCCCAGCACTTTGGGAGCTTAAGTGGGTAGATTACTTGAGCCCAGGAATTCAGACCA GCTTGGCACATGGTGAACCTGTTCTATAAAAATAATCTGGCTTTGAGCATATGCCTGTGGTC CAGCACTGAGAGGCTAGTGAAGATTGCTGAGCCCAGAGCCAAAGGTTGCAGTGAGCAAGTCA AGGCAAAATTTTGACAGGGAAGGAAGTAACTGCAAAACCACTAGGCTTTAGTAGGTACTTAT GATGTGGATTGCTGGTGTCCAGCATGACCCATAAACAGGTCAGAAGAATGATGGAATGTTTT AGAATAAACTCCTGCTTATAGTATACTACACAGTTCAAAAGATGTTTAAAATGCTTTTGTAT TTACTGCCATGTAATTGAAATATATAGATTATTGTAACCTTTCAACCTGAAAATCAAGCAGT ATGAGAGTTTAGTTATTTGTATGTGTCACTAGTGTCTAATGAAGCTTTTAAAATCTACAATT TCTTCTTTAAAAATATTTATTAATGTGAATGGAATATAACAATTCAGCTTAATTCCCCAACC TTATTCTGTGTGTAGACATTGTATTCCACAATTTTGAATGGCTGTGTTTTACCTCTAAATAA ATGAATTCAGAGAAAAAAAAAAAAAAA

FIGURE 41

MGSVLGLCSMASWIPCLCGSAPCLLCRCCPSGNNSTVTRLIYALFLLVGVCVACVMLIPGME
EQLNKIPGFCENEKGVVPCNILVGYKAVYRLCFGLAMFYLLLSLLMIKVKSSSDPRAAVHNG
FWFFKFAAAIAIIIGAFFIPEGTFTTVWFYVGMAGAFCFILIQLVLLIDFAHSWNESWVEKM
EEGNSRCWYAALLSATALNYLLSLVAIVLFFVYYTHPASCSENKAFISVNMLLCVGASVMSI
LPKIQESQPRSGLLQSSVITVYTMYLTWSAMTNEPETNCNPSLLSIIGYNTTSTVPKEGQSV
QWWHAQGIIGLILFLLCVFYSSIRTSNNSQVNKLTLTSDESTLIEDGGARSDGSLEDGDDVH
RAVDNERDGVTYSYSFFHFMLFLASLYIMMTLTNWSRYEPSREMKSQWTAVWVKISSSWIGI
VLYVWTLVAPLVLTNRDFD

FIGURE 42

GCGAGAAGAAGCTGTCTCCATCTTGTCTGTATCCCGCTGCTTCTTGNGACGTTGTGGAGAT
GGGGAGCGTCCCTGGGGCTGTGCTCCATGGCGAGCTGGATACCATGTTTGTGTGGAAGTGCC
CCGTGTTTGCTATGCCGATGCTGTCCTAGTGGAAACAANTCCACTGTAACTAGATTGATCTA
TGCACTTTTCTTGCTTGTTGGAGTATGTGTAGCTTGTGAATGTTGATACCAGGAATGGAAG
AACAACTGAATAAGATTCCTGGATTTTGTGAGAATGAGAAAGGTGTTGTCCCTTGTAACATT
TTGGTTGGCTATAAAGCTGTATATCGTTTGTGCTTTGGTTTGGCTATGTTCTATCTTCTT
CTCTTTACTAATGATCAAAGTGAAGAGTAGCAGTGATCCTAGAGCTGCACAATGGAT
TTTGGTTCTTTAAATTTGCTGCAGCAATTGCAATTATTATTGGGGC

FIGURE 43

GTTATTGTGAACTTTGTGGAGATGGGAGGTCNTGGGGCTGTGTTCCATGGCGAGCTGGATAC
CANGTTTGTGTGGAAGTGCCCCGTGTTTGNTATGCCGATGCTGTCCTAGTGGAAACAANTCC
ACTGTAATTAGATTGATNTATGCACTTTTNTTGCTTGTTGGAGTANGTGTAGCTTGTGTAAT
GTTGATACCAGGAATGGAAGAACAACTGAATAAGATTCCTGGATTTTGTGAGAATGAGAAAG
GTGTTGTCCCTTGTAACATTTTGGTTGGCTATAAAGCTGTATATNGTTTGTGCTTTGGTTTG
GCTANGTTCTATNTTCTTCTCTCTTTTACTAATGATCAAAGTGAAGAGTAGCAGTGATCCTAG
AGCTGCAGTGCACAATGGATTTTGGTTTTTAAATTTGCTGCAGCAATTGCAATTATTATTG
GGGC

FIGURE 44

FIGURE 45

FIGURE 46

GTCGTTTCCAGCCAAGTGGACCTGATCGATGGCCCTCCTGAATTTATCACGATATTTGATTTATTAGCGATGCC $\tt CCCTGGTTTGTGTTACGCACACACGTGCACACAAGGCTCTGGCTCGCTTCCCTCGTTTCCAGCTCC$ TGGGCGAATCCCACATCTGTTTCAACTCTCCGCCGAGGGCGAGCAGGAGCGAGAGTGTGTCGAATCTGCGAGTG AAGAGGGACGAGGGAAAAGAAACAAAGCCACAGACGCAACTTGAGACTCCCGCATCCCAAAAGAAGCACCAGAT $\texttt{CAGCAAAAAAAGAAG} \underbrace{\textbf{ATG}} \texttt{GGCCCCCGAGCCTCGTGCTGTTGCTTGCTGTCGCAACTGTTTTCCCTGCTGGG}$ TGGAAGCTCGGCCTTCCTGTCGCACCACCGCCTGAAAGGCAGGTTTCAGAGGGACCGCAGGAACATCCGCCCCA ACATCATCCTGGTGCTGACGGACGACCAGGATGTGGAGCTGGGTTCCATGCAGGTGATGAACAAGACCCGGCGC ATCATGGAGCAGGGCGGGCGCACTTCATCAACGCCTTCGTGACCACACCCATGTGCTGCCCCTCACGCTCCTC CATCCTCACTGGCAAGTACGTCCACAACCACAACACCTACACCAACAATGAGAACTGCTCCTCGCCCTCCTGGC AGGCACAGCACGAGACCCCTTTGCCGTGTACCTCAATAGCACTGGCTACCGGACAGCTTTCTTCGGGAAG TATCTTAATGAATACAACGGCTCCTACGTGCCACCCGGCTGGAAGGAGTGGGTCGGACTCCTTAAAAACTCCCG CTTTTATAACTACACGCTGTGTCGGAACGGGGTGAAAGAGAAGCACGGCTCCGACTACTCCAAGGATTACCTCA CAGACCTCATCACCAATGACAGCGTGAGCTTCTTCCGCACGTCCAAGAAGATGTACCCGCACAGGCCAGTCCTC ATGGTCATCAGCCATGCAGCCCCCACGGCCCTGAGGATTCAGCCCCACAATATTCACGCCTCTTCCCAAACGC ATCTCAGCACATCACGCCGAGCTACAACTACGCGCCCAACCCGGACAAACACTGGATCATGCGCTACACGGGGC CCATGAAGCCCATCCACATGGAATTCACCAACATGCTCCAGCGGAAGCGCTTGCAGACCCTCATGTCGGTGGAC GACTCCATGGAGACGATTTACAACATGCTGGTTGAGACGGGCGAGCTGGACAACACGTACATCGTATACACCGC CGACCACGGTTACCACATCGGCCAGTTTGGCCTGGTGAAAGGGAAATCCATGCCATATGAGTTTGACATCAGGG TCCCGTTCTACGTGAGGGGCCCCAACGTGGAAGCCGGCTGTCTGAATCCCCACATCGTCCTCAACATTGACCTG GGACACGGAGCGGCCGGTGAATCGGTTTCACTTGAAAAAGAAGATGAGGGTCTGGCGGGACTCCTTCTTGGTGG AGAGAGCCAAGCTGCTACACAAGAGAGACAATGACAAGGTGGACGCCCAGGAGGAGAACTTTCTGCCCAAGTAC CAGCGTGTGAAGGACCTGTGTCAGCGTGCTGAGTACCAGACGGCGTGTGAGCAGCTGGGACAGAAGTGGCAGTG TGTGGAGGACGCCACGGGGAAGCTGAAGCTGCATAAGTGCAAGGGCCCCATGCGGCTGGGCGGCAGCAGAGCCC TCTCCAACCTCGTGCCCAAGTACTACGGGCAGGGCAGGCCTGCACCTGTGACAGCGGGGGACTACAAGCTC AGCCTGGCCGGACGCCGGAAAAAACTCTTCAAGAAGAAGTACAAGGCCAGCTATGTCCGCAGTCGCTCCATCCG CTCAGTGGCCATCGAGGTGGACGGCAGGGTGTACCACGTAGGCCTGGGTGATGCCGCCAGCCCCGAAACCTCA CCAAGCGGCACTGGCCAGGGGCCCCTGAGGACCAAGATGACAAGGATGGTGGGGACTTCAGTGGCACTGGAGGC CTTCCCGACTACTCAGCCGCCAACCCCATTAAAGTGACACATCGGTGCTACATCCTAGAGAACGACACAGTCCA GTGTGACCTGGACCTGTACAAGTCCCTGCAGGCCTGGAAAGACCACAAGCTGCACATCGACCACGAGATTGAAA CCCTGCAGAACAAATTAAGAACCTGAGGGAAGTCCGAGGTCACCTGAAGAAAAAGCGGCCAGAAGAATGTGAC TGTCACAAAATCAGCTACCACACCCAGCACAAAGGCCGCCTCAAGCACAGAGGCTCCAGTCTGCATCCTTTCAG GAAGGGCCTGCAAGAAGGACAAGGTGTGGCTGTTGCGGGAGCAGAAGCGCAAGAAGAAACTCCGCAAGCTGC TCAAGCGCCTGCAGAACAACGACACGTGCAGCATGCCAGGCCTCACGTGCTTCACCCACGACAACCAGCACTGG ${\tt CAGACGGCGCCTTTCTGGGCCCTTTCTGTGCCTGCACCAGCGCCAACAATAACACGTACTGGTGCAT}$ GAGGACCATCAATGAGACTCACAATTTCCTCTTCTGTGAATTTGCAACTGGCTTCCTAGAGTACTTTGATCTCA ACACAGACCCCTACCAGCTGATGAATGCAGTGAACACACTGGACAGGGATGTCCTCAACCAGCTACAGGCACAGCAGCTACAG AAGCTATGAGCAATACAGGCAGTTTCAGCGTCGAAAGTGGCCAGAAATGAAGAGACCTTCTTCCAAATCACTGG GACAACTGTGGGAAGGCTGGGAAGGT**TAA**GAAACAACAGAGGTGGACCTCCAAAAACATAGAGGCATCACCTGA CTGCACAGGCAATGAAAAACCATGTGGGTGATTTCCAGCAGACCTGTGCTATTGGCCAGGAGGCCTGAGAAAGC AAGCACGCACTCTCAGTCAACATGACAGATTCTGGAGGATAACCAGCAGGAGGAGAGATAACTTCAGGAAGTCC ATTTTTGCCCCTGCTTTTGCTTTGGATTATACCTCACCAGCTGCACAAAATGCATTTTTTCGTATCAAAAAGTC ${ t TCCCAAGGGCGAAAGTCATTGGAATTTTTAAATCATAGGGGAAAAGCAGTCCTGTTCTAAATCCTCTTATTCTT$ TTGGTTTGTCACAAAGAAGGAACTAAGAAGCAGGACAGAGGCAACGTGGAGAGGCTGAAAACAGTGCAGAGACG TTTGACAATGAGTCAGTAGCACAAAAGAGATGACATTTACCTAGCACTATAAACCCTGGTTGCCTCTGAAGAAA CTGCCTTCATTGTATATGTGACTATTTACATGTAATCAACATGGGAACTTTTAGGGGAACCTAATAAGAAAT CCCAATTTCAGGAGTGGTGGTCAATAAACGCTCTGTGGCCAGTGTAAAAGAAAAA

FIGURE 47

MGPPSLVLCLLSATVFSLLGGSSAFLSHHRLKGRFQRDRRNIRPNIILVLTDDQDVELGSMQ
VMNKTRRIMEQGGAHFINAFVTTPMCCPSRSSILTGKYVHNHNTYTNNENCSSPSWQAQHES
RTFAVYLNSTGYRTAFFGKYLNEYNGSYVPPGWKEWVGLLKNSRFYNYTLCRNGVKEKHGSD
YSKDYLTDLITNDSVSFFRTSKKMYPHRPVLMVISHAAPHGPEDSAPQYSRLFPNASQHITP
SYNYAPNPDKHWIMRYTGPMKPIHMEFTNMLQRKRLQTLMSVDDSMETIYNMLVETGELDNT
YIVYTADHGYHIGQFGLVKGKSMPYEFDIRVPFYVRGPNVEAGCLNPHIVLNIDLAPTILDI
AGLDIPADMDGKSILKLLDTERPVNRFHLKKKMRVWRDSFLVERGKLLHKRDNDKVDAQEEN
FLPKYQRVKDLCQRAEYQTACEQLGQKWQCVEDATGKLKLHKCKGPMRLGGSRALSNLVPKY
YGQGSEACTCDSGDYKLSLAGRRKKLFKKKYKASYVRSRSIRSVAIEVDGRVYHVGLGDAAQ
PRNLTKRHWPGAPEDQDDKDGGDFSGTGGLPDYSAANPIKVTHRCYILENDTVQCDLDLYKS
LQAWKDHKLHIDHEIETLQNKIKNLREVRGHLKKKRPEECDCHKISYHTQHKGRLKHRGSSL
HPFRKGLQEKDKVWLLREQKRKKKLRKLLKRLQNNDTCSMPGLTCFTHDNQHWQTAPFWTLG
PFCACTSANNNTYWCMRTINETHNFLFCEFATGFLEYFDLNTDPYQLMNAVNTLDRDVLNQL
HVQLMELRSCKGYKQCNPRTRNMDLDGGSYEQYRQFQRRKWPEMKRPSSKSLGQLWEGWEG

FIGURE 48

AACAAAGTTCAGTGACTGAGGGGCTGAGCGGAGGCTGCTGAAGGGGAGAAAGGAGTGAGGA ${\tt GCTGCTGGGCAGAGAGGGACTGTCCGGCTCCCAG}$ TCGTGGGATGATCACAGGTGCTGCTGTGGCGGTCCTGCTGCTGCTGCTGCTGCCACC $\tt TGCCTTTTCCACGGACGGCAGGACTGTGACGTGGAGGAGGAACCGTACAGCTGCAGGGGGAAA$ CCGAGTCCGCCGGCCCAGCCTTGGCCCTTCCGGCGGGGGCCACCTGGGAATCTTTCACC ATCACCGTCATCCTGGCCACGTATCTCATGTGCCGAATGTGGGCCTCCACCACCACCACCAC CCCCGCCACACCCTCACCACCTCCACCACCACCACCCCCACCGCCACCACCACCCCCACCGCCA TTCCCAGGACAAGTGGACCCCATGTTTCCATGTGGAAGGATGCATCTCTGGGGTGAACGAGG GGAACAATAGACTGGGGCTTGCTCCAGCTGCATTTGCATGGCATGCCCCAGTGTACTATGGC AGCAGAGATGGAGGAACACTGGGTCTGCAGTGCTGAAGGGTTTGGGGAGTGGAGAGCAAGG GTGCTCTTTCGGGGCTGGACAGCCCGTCTTGTGACAGTGACTCCCAGTGAGCCCCAGAAATG ACAAGCGTGTCTTGGCAGAGCCAGCACACAGTGGATGTGAAGTGCCCGTCTTGACCTCCTC ATCAGGCTGCTGCAGGCCTCTGGCGGCAGGGCACTGGGAGAGGCCCTGAGAATGTCCTTTT GGTTTGGAGAAGGCAGTGTGAGGCTGCACAGTCAATTCATCGGTGCCTTAGTCCAAGAAAAT

FIGURE 49

MLGLLGSTALVGWITGAAVAVLLLLLLLATCLFHGRQDCDVERNRTAAGGNRVRRAQPWPFR RRGHLGIFHHHRHPGHVSHVPNVGLHHHHHPRHTPHHLHHHHHPRHAR

FIGURE 50

 ${\tt GGCGGCTGCTGAGCTGCTTGAGGTGCAGTGTTGGGGGATCCAGAGCC} {\color{red} {\bf ATC}} {\tt TCGGACCTGCTA}$ CTACTGGGCCTGATTGGGGGCCTGACTCTTTACTGCTGACGCTGCTGGCCTTTGCCGG GTACTCAGGGCTACTGGGTGGAAGTGAGTGCTGGGTCACCCCCCATCCGCAACGTCA CTGTGGCCTACAAGTTCCACATGGGGGCTCTATGGTGAGACTGGGCGGCTTTTCACTGAGAGC TGCAGCATCTCTCCCAAGCTCCGCTCCATCGCTGTCTACTATGACAACCCCCACATGGTGCC CCCTGATAAGTGCCGATGTGCCGTGGGCAGCATCCTGAGTGAAGGTGAGGAATCGCCCTCCC CTGAGCTCATCGACCTCTACCAGAAATTTGGCTTCAAGGTGTTCTCCTTCCCGGCACCCAGC TGTCCATCCTGCCTTGGACACCTACATCAAGGAGCGGAAGCTGTGTGCCTATCCTCGGCTGG AGATCTACCAGGAAGACCAGATCCATTTCATGTGCCCACTGGCACGGCAGGGAGACTTCTAT GTGCCTGAGATGAAGGAGACAGAGTGGAAATGGCGGGGGCTTGTGGAGGCCATTGACACCCA GGTGGATGGCACAGGAGCTGACACAATGAGTGACACGAGTTCTGTAAGCTTGGAAGTGAGCC CTGGCAGCCGGGAGACTTCAGCTGCCACACTGTCACCTGGGGCGAGCAGCCGTGGCTGGGAT GACGGTGACACCCGCAGCGAGCACAGCTACAGCGAGTCAGGTGCCAGCGGCTCCTCTTTTGA GGAGCTGGACTTGGAGGGCCCTTAGGGGAGTCACGGCTGGACCCTGGGACTGAGC $\verb|CCCTGGGGACTACCAAGTGGCTCTGGGAGCCCACTGCCCCTGAGAAGGGCCAAGGAGTAACCC| \\$ ATGGCCTGCACCCTCCTGCAGTGCAGTTGCTGAGGAACTGAGCAGACTCTCCAGCAGACTCT CCAGCCCTCTTCCTCCTGGGGGGGGGGGGGGTTCCTGAGGGACCTGACTTCCCCTGC TCCAGGCCTCTTGCTAAGCCTTCTCCTCACTGCCCTTTAGGCTCCCAGGGCCAGAGGAGCCA GGGACTATTTTCTGCACCAGCCCCCAGGGCTGCCGCCCCTGTTGTGTCTTTTTTTCAGACTC ACAGTGGAGCTTCCAGGACCCAGAATAAAGCCAATGATTTACTTGTTTCACCTGGAAAAAAA AAAAAAAAA

FIGURE 51

MSDLLLLGLIGGLTLLLLTLLAFAGYSGLLAGVEVSAGSPPIRNVTVAYKFHMGLYGETGR LFTESCSISPKLRSIAVYYDNPHMVPPDKCRCAVGSILSEGEESPSPELIDLYQKFGFKVFS FPAPSHVVTATFPYTTILSIWLATRRVHPALDTYIKERKLCAYPRLEIYQEDQIHFMCPLAR QGDFYVPEMKETEWKWRGLVEAIDTQVDGTGADTMSDTSSVSLEVSPGSRETSAATLSPGAS SRGWDDGDTRSEHSYSESGASGSSFEELDLEGEGPLGESRLDPGTEPLGTTKWLWEPTAPEK GKE

FIGURE 52

FIGURE 53

MTLRPSLLPLHLLLLLLSAAVCRAEAGLETESPVRTLQVETLVEPPEPCAEPAAFGDTLHI HYTGSLVDGRIIDTSLTRDPLVIELGQKQVIPGLEQSLLDMCVGEKRRAIIPSHLAYGKRGF PPSVPADAVVQYDVELIALIRANYWLKLVKGILPLVGMAMVPALLGLIGYHLYRKANRPKVS KKKLKEEKRNKSKKK

FIGURE 54

FIGURE 55

CCGAAAGTCCCGTCCGGACCCTCCAAGTGGAGACCCTGGTGGAGCCCCCAGAACCATGTGCC
GAGCCCGCTGCTTTTGGAGACACGCTTCACATACACTACACGGGAAGCTTGGTAGATGGACG
TATTATTGACACCTCCCTGACCAGAGACCCTCTGGTTATAGAACTTGGCCAAAAGCAGGTGA
TTCCAGGTCTGGAGCAGAGTCTTCTCGACATGTGTGTGGGAGAAGCGAAGGGCAATCATT
CCTTCTCACTTGGCCTATGGAAAACGGGGATTTCCACCATCTGTCCCAGCGGATGCAGTGGT
GCAGTATGACGTGGAGCTGATTGCACTAATCCGAGCCAACTACTGGCTAAAGCTGGTGAAGG
GCATTTTGCCTCTGGTAGGGATGGCCATGGTGCCAGCCCTCCTGGGCCTCATTGGGTATCAC
CTATACAGAAAGGCCAATAGACCCAAAGTCTCCAAAAAGAAGCTCAAGGAAGAAACGAAA
CAAGAGCAAAAAGAAATAATAAATAATAATTTTTAAAAAACTTTAAAA

FIGURE 56

CTGCTGCATCCGGGTGTCTGGAGGCTGTGGCCGTTTTGTTTTCTTGGCTAAAATCGGGGGAG TGAGGCGGGCCGGCGCGCGACACCGGGCTCCGGAACCACTGCACGACGGGGCTGGACTG ACCTGAAAAAA**ATG**TCTGGATTTCTAGAGGGCTTGAGATGCTCAGAATGCATTGACTGGGGG GAAAAGCGCAATACTATTGCTTCCATTGCTGCTGGTGTACTATTTTTTTACAGGCTGGTGGAT TATCATAGATGCAGCTGTTATTTATCCCACCATGAAAGATTTCAACCACTCATACCATGCCT GTGGTGTTATAGCAACCATAGCCTTCCTAATGATTAATGCAGTATCGAATGGACAAGTCCGA GGTGATAGTTACAGTGAAGGTTGTCTGGGTCAAACAGGTGCTCGCATTTGGCTTTTCGTTGG TTTCATGTTGGCCTTTGGATCTCTGATTGCATCTATGTGGATTCTTTTTTGGAGGTTATGTTG CTAAAGAAAAAGACATAGTATACCCTGGAATTGCTGTATTTTTCCAGAATGCCTTCATCTTT TTTGGAGGGCTGGTTTTTAAGTTTGGCCGCACTGAAGACTTATGGCAG**TGA**ACACATCTGAT TTCCCACAGCACAACAGCCCTGCATGGGTTTGTTTGTTTTTTTACTGCTCACTCCCAACCTT TTGTAATGCCATTTTCTAAACTTATTTCTGAGTGTAGTCTCAGCTTAAAGTTGTGTAATACT AAAATCACGAGAACACCTAAACAACAACAAAATCTATTGTGGTATGCACTTGATTAACTT ATAAAATGTTAGAGGAAACTTTCACATGAATAATTTTTGTCAAATTTTATCATGGTATAATT TGTAAAAATAAAAAGAAATTACAAAAGAAATTATGGATTTGTCAATGTAAGTATTTGTCATA TCTGAGGTCCAAAACCACAATGAAAGTGCTCTGAAGATTTAATGTGTTTATTCAAATGTGGT CTCTTCTGTGTCAAATGTTAAATGAAATATAAACATTTTTTTAGTTTTTTAAAATATTCCGTGG TCAAAATTCTTCCTCACTATAATTGGTATTTACTTTTACCAAAAATTCTGTGAACATGTAAT GTAACTGGCTTTTGAGGGTCTCCCAAGGGGTGAGTGGACGTGTTGGAAGAGAAGCACCAT GGTCCAGCCACCAGGCTCCCTGTGTCCCTTCCATGGGAAGGTCTTCCGCTGTGCCTCTCATT CCAAGGCAGGAAGATGTGACTCAGCCATGACACGTGGTTCTGGTGGGATGCACAGTCACTC CACATCCACCACTG

FIGURE 57

MSGFLEGLRCSECIDWGEKRNTIASIAAGVLFFTGWWIIIDAAVIYPTMKDFNHSYHACGVI ATIAFLMINAVSNGQVRGDSYSEGCLGQTGARIWLFVGFMLAFGSLIASMWILFGGYVAKEK DIVYPGIAVFFQNAFIFFGGLVFKFGRTEDLWQ

FIGURE 58

FIGURE 59

TGGACGGACCTGAAAAAATGTTTGGATTTNTAGAGGGNTTGAGATGTTCAGAATGCATGAC
TGGGGGAAAAGCGCAAATACTATTGCTTCCATTGCTGCTGGTGTANTATTTTTTTACAGGCTG
GTGGATTATCATAGATGCAGNTGTTATTTTATCCCACCATGAAAGATTTCAACCANTCATACC
ATGCCTGTGGTGTTATAGCAACCATAGCCTTCNTAATGATTAATGCAGTATCGAATGGACAA
GTCCGAGGTGATAGTTACAGTGAAGGTTGTTTGGGTCAAACAGGTGCTCGCATTTGGCTTTT
CGTTGGTTTCATGTTGGCCTTTGGATCTCTGATTGCATCTATGTGGATTCTTTTTGGAGGTT
ATGTTGCTAAAGAAAAAGACATAGTATACCCTGGAATTGNTGTATTTTTCCAGAATGCCTTC
ATCTTTTTTGGAGGGCTGGTTTTTAAGTTTGGCCGCACTGAAGANTTATGGCAGTG

FIGURE 60

GGACACCGGGTTCCGGACCAATGCANGACGGGGTGGANTGACCTGAAAAAAATGTTTGGATT
TTTAGAGGGCTTGAGATGNTCAGAATGCATTGACTGGGGGAAAAGCGCAATANTATTGCTTT
CCATTGCTGCTGGTGTACTATTTTTTACAGGGTTGGTGGATTATCATAGATGCAGCTGTTATT
TATCCCACCATGAAAGATTTNAACCACTCATACCATGCCTGTGGTGTTATAGCAACCATAGC
CTTCCTAATGATTAATGCAGTATCGAATGGACAAGTCCGAGGTGATAGTTACAGTGAAGGTT
GTTTGGGTCAAACAGGTGNTCGCATTTGGCTTTTCGTTGGTTTCATGTTGGCCTTTTGGATTT
CTGATTGNATTCTATGCGGATTCTTCTTGGAGGTTATGTTGCTAAAGAAAAAGACATAGTAT
ACCCTGGAATTNCTNTATTTTTCCAGAATGCC

FIGURE 61

TAGAGGGCTTGAGATGCTCAGAATGCATTGACTGGGGGGAAAAGCGCAATANTATTGCTTCC
ATTGNTGNTGGTGTANTATTTTTTTACAGGCTGGTGGATTATNATAGATGCAGCTGTTATTT
ATCCCACCATGAAAGATTTNAACCANTCATACCATGCCTGTGGTGTTATAGCAACCATAGCC
TTCCTAATGATTAATGCAGTATNGAATGGACAAGTCCGAGGTGATAGTTACAGTGAAGGTTG
TTTGGGTCAAACAGGTGNTNGCATTTGGCTTTTNGTTGGTTTCATGTTGGCCTTTTGGATCTN
TGATTGCATTTATGTGGATTNTTTTTTGGAGGTTATGTTGCTAAAGNAAAAGACATAGTATAC
CCTGT

FIGURE 62

FIGURE 63

CGACGCCGGCGTG**ATG**TGGCTTCCGCTGGTGCTGCTGCTGCTGCTGCCGTCC TCTGCAAAGTTTACTTGGGACTATTCTCTGGCAGCTCCCCGAATCCTTTCTCCGAAGATGTC AAACGGCCCCCAGCGCCCCTGGTAACTGACAAGGAGGCCAGGAAGAAGGTTCTCAAACAAGC AAGGGCAGCTGGACTGGGCTCCCCTGTCCTCTCTCTTTTTGACATCATGGTACTGGAAGGGCCC
AATGGCCGAAAGGAGTACCCCATGTACAGTGGAGAAAGCCTACATTCAGGGCCTCAAGGA
GAAGTTTCCACAGGAGGAAGCTATCATTGACAAGTATATAAAGCTGGTTAAGGTGGTATCCA
GTGGAGCCCCTCATGCCATCCTGTTGAAATTCCTCCCATTGCCCTGGTTCAGCTCCTCGAC
AGGTGTGGGCTGCTGACTCCTTCTCCATTCCTTCAAGCATCCACCAAGAGCCTGGCTGA
GGTCCTGCAGCAGCAGCTCCTCTGAGCTCCATGCCCCTGCTGGTCAACCACTAC
CTTACGGTGTCACCCCCAACCACAGTGCCTTTTCAACACTCCAACACCACTAC CTGTTTTCATCTGCCTGCGAGGCACCAAGGAAGACCTGCATCTGCCGTCCACCAACTACTAT GTTTACTATGACACGGACATGGACCAGGCGATGGAGCGCTACGTCCATGCCAGGGAAGA GGCTGCGGAACACATCCCTCTTCTTCTTCTTCTCCCATCCAGCCAAAGATCCGACCTGGG AGGACCGATTCCCAGGCCGTCCACCATGATCATGCTCATACCCACTGCCTACGAGTGGTTT GAGGAGTGGCAGGCGGAGCTGAAGGGAAAGCGGGGCAGTGACTATGAGACCTTCAAAAACTC CTTTGTGGAAGCCTCTATGTCAGTGGTCCTGAAACTGTTCCCACAGCTGGAGGGGAAGGTGG CTTTGTGGAAGCCTCTATGTCAGTGGTCCTGAAACTGTTCCCACAGCTGGAGGGGAAGGTGG
AGAGTGTGACTGCAGGATCCCCACTCACCAACCAGTTCTATCTGGCTGCTCCCCGAGGTGCC
TGCTACGGGGCTGACCATGACCTGGGCCGCCTGCACCCTTGTGTGATGGCCTCCTTGAGGGC
CCAGAGCCCCATCCCCAACCTCTATCTGACAGGCCAGGATATCTTCACCTGTGGACTGGTCG
GGGCCCTGCAAGGTGCCCTGCTGCAGCAGCAGCCCATCCTGAAGCGGAACTTGTACTCAGAC
CTTAAGAATCTTGATTCTAGGATCCGGGCACAGAAGAAAAAGAATTAGTTCCATCAGGGAGG
AGTCAGAGGAATTTGCCCAATGGCTGGGGCATCTCCCTTGACTTACCCATAATGTCTTTCTG
CATTAAGTTCCTTGCACGTATAAAGCACTCTAATTTGGTTCTGATGCCTGAAGAGAGGGCCTAG
TTTAAATCACAATTCCGAATCTGGGGCAATGGAATCACTGCTTCCAGCTGGGGCAGGTGAGA
TCTTTACGCCTTTTATAACATGCCATCCCTACTAATAGGATATTTGACTTGGATAGCTTCAATG
TCTCATGACGAGCGGCGCTCTGCATCCCTCACCCATGCCTCCAACCTCAGTGATCAAAGCGA
ATATTCCATCTGTGGATAGAACCCCCTGGCAGTTCTCAGCTCACCTGGTGGTTCAGTTC ${ t ATATTCCATCTGTGGATAGAACCCCTGGCAGTGTTGTCAGCTCAACCTGGTGGGTTCAGTTC}$ TGGACTGGTAAGATGAATACTTTGCTGGGCTGAAGCAGGCTGCAGGGCATTCCAGCCAAGGG CACAGCAGGGGACAGTGCAGGGAGGTGTGGGGTAAGGGAAGTCACATCAGAAAAGGGA GTATCAGACATACGAAAGGTCTCTTTGTAGTTCGTGTTAATGTAACATTAATAAATTTATTG ATTCCATTGCTTTAAAAAAAAAAAAAAA

FIGURE 64

MWLPLVLLLAVLLLAVLCKVYLGLFSGSSPNPFSEDVKRPPAPLVTDKEARKKVLKQAFSAN QVPEKLDVVVIGSGFGGLAAAAILAKAGKRVLVLEQHTKAGGCCHTFGKNGLEFDTGIHYIG RMEEGSIGRFILDQITEGQLDWAPLSSPFDIMVLEGPNGRKEYPMYSGEKAYIQGLKEKFPQ EEAIIDKYIKLVKVVSSGAPHAILLKFLPLPVVQLLDRCGLLTRFSPFLQASTQSLAEVLQQ LGASSELQAVLSYIFPTYGVTPNHSAFSMHALLVNHYMKGGFYPRGGSSEIAFHTIPVIQRA GGAVLTKATVQSVLLDSAGKACGVSVKKGHELVNIYCPIVVSNAGLFNTYEHLLPGNARCLP GVKQQLGTVRPGLGMTSVFICLRGTKEDLHLPSTNYYVYYDTDMDQAMERYVSMPREEAAEH IPLLFFAFPSAKDPTWEDRFPGRSTMIMLIPTAYEWFEEWQAELKGKRGSDYETFKNSFVEA SMSVVLKLFPQLEGKVESVTAGSPLTNQFYLAAPRGACYGADHDLGRLHPCVMASLRAQSPI PNLYLTGQDIFTCGLVGALQGALLCSSAILKRNLYSDLKNLDSRIRAQKKKN

FIGURE 65

GCAGCGGCGAGGCGGTGGTGGCTGAGTCCGTGGTGGCAGAGGCGAAGGCGACAGCTCTA $\tt GGGGTTGGCACCGGCCCCGAGAGGAGGAGGAGGAGGTCCGGATAGGGCTGACGCTGCTGTG$ TGCGGTGCTGAGCTTGGCCTCGGCGTCCTCGGATGAAGAAGGCAGCCAGGATGAATCCT GTAGTTGCTGGTCAAATATTTCTTGATTCAGAAGAATCTGAATTAGAATCCTCTATTCAAGA AGAGGAAGACAGCCTCAAGAGCCAAGAGGGGGGAAAGTGTCACAGAAGATATCAGCTTTCTAG ACCGCCATTGAAGGCACAGGGGGGGGGCCCTGCCACTTCCCTTTTCCTTAGATAA GGAGTATGATGAATGTACATCAGATGGGAGGGAAGATGGCAGACTGTGGTGTGCTACAACCT ATGACTACAAAGCAGATGAAAAGTGGGGCTTTTTGTGAAACTGAAGAAGAGGCTGCTAAGAGA CGGCAGATGCAGGAAGCAGAAATGATGTATCAAACTGGAATGAAAATCCTTAATGGAAGCAA CCAAAGCCCTGGAGAGAGTGTCATATGCTCTTTTATTTGGTGATTACTTGCCACAGAATATC CAGGCAGCGAGAGATGTTTGAGAAGCTGACTGAGGAAGGCTCTCCCAAGGGACAGACTGC TCTTGGCTTTCTGTATGCCTCTGGACTTGGTGTTAATTCAAGTCAGGCAAAGGCTCTTGTAT ATTATACATTTGGAGCTCTTGGGGGCAATCTAATAGCCCACATGGTTTTGGTAAGTAGACTT TAGTGGAAGGCTAATAATATTAACATCAGAAGAATTTGTGGTTTATAGCGGCCACAACTTTT ATTCTTGTTAATGGATATAACACATGGAATCTACATGTAAATGAAAGTTGGTGGAGTCCACA ATTTTTCTTTAAAATGATTAGTTTGGCTGATTGCCCCTAAAAAGAGAGATCTGATAAATGGC TCTTTTTAAATTTTCTCTGAGTTGGAATTGTCAGAATCATTTTTTACATTAGATTATCATAA TTTTAAAAATTTTTCTTTAGTTTTTCAAAATTTTGTAAATGGTGGCTATAGAAAAACAACAT GAAATATTATACAATATTTTGCAACAATGCCCTAAGAATTGTTAAAATTCATGGAGTTATTT CCATTTATTCTGGTCATTTATTGCTAGTGACACTGTGCCTGCTTCCAGTAGTCTCATTTTCC CTATTTTGCTAATTTGTTACTTTTTCTTTGCTAATTTGGAAGATTAACTCATTTTTAATAAA

FIGURE 66

MRVRIGLTLLLCAVLLSLASASSDEEGSQDESLDSKTTLTSDESVKDHTTAGRVVAGQIFLD SEESELESSIQEEEDSLKSQEGESVTEDISFLESPNPENKDYEEPKKVRKPALTAIEGTAHG EPCHFPFLFLDKEYDECTSDGREDGRLWCATTYDYKADEKWGFCETEEEAAKRRQMQEAEMM YQTGMKILNGSNKKSQKREAYRYLQKAASMNHTKALERVSYALLFGDYLPQNIQAAREMFEK LTEEGSPKGQTALGFLYASGLGVNSSQAKALVYYTFGALGGNLIAHMVLVSRL

FIGURE 67

FIGURE 68

MACRCLSFLLMGTFLSVSQTVLAQLDALLVFPGQVAQLSCTLSPQHVTIRDYGVSWYQQRAG SAPRYLLYYRSEEDHHRPADIPDRFSAAKDEAHNACVLTISPVQPEDDADYYCSVGYGFSP

FIGURE 69

GCCGCCCGCCCGAGACCGGGCCCGGGGGCGCGGGGCGGGGTGCGGCGCCCGGGGCGC CGATGACCGCGGAGCGCACGCCGGGGCCCGGCCTGACCCCGCCGCCGCCGGCTGAGCCC CCCGCCGAGGTCCGGACAGGCCGAG**ATG**ACGCCGAGCCCCCTGTTGCTGCTGCTGCCGC CGCTGCTGCTGGGGGCCTTCCCACC<mark>GGC</mark>CGCCGCCCGAGGCCCCCCAAAGATGGCGGAC AAGGTGGTCCCACGGCAGGTGGCCCGGCTGGGCCGCACTGTGCGGCTGCAGTGCA GGGGGACCCGCCGCCGCTGACCATGTGGACCAAGGATGGCCGCACCATCCACAGCGGCTGGA GCCGCTTCCGCGTGCTGCCGCAGGGCTGAAGGTGAAGCAGCATCCACAGCGGCTGAGCTTCCGCGTGCTGCCGCCGCCGCCTGAGCGTGAAGCAGCCTCGTCGTCGTACGTGTGCAAGCCAACCCCTCGTCGTCGTGGATGACATTAGCCCAGGGAAGGAGAGCCTGGGGCCCGACAGCTCCTCTGGGGGTCAAG GCGGACCCGTTCCAAGCCCGTGCTCACAGGCACGCACCCCGTGAACACGACGGTGGACTTCG GGGGGACCCGTTCCAAGCCCGTGCTCACAGGCACCCCGTGAACACGACGGTGACTTCG GGGGGACCACGTCCTTCCAGTGCAAGGTGCGCAGACGTGAAGCCGGTGATCCAGTGGCTG AAGCGCGTGGAGTACGGCGCCGAGGGCCGCCACAACTCCACCATCGATGTGGGCGGCCAGAA GTTTGTGGTGCTGCCCACGGGTGACGTGTGGTCGCGGCCCGACGCTCCTACCTCAATAAGC TGCTCATCACCCGTGCCCCCAGGACGATGCGGCATGTACATCTGCCTTGGCGCCAACACC ATGGGCTACAGCTTCCGCAGCGCCTTCCTCACCGTGCTGCCAGACCCAAAACCGCCAGGGCC ACACTCACACGTGGAGGGCAAGGTCCACCAGCACATCCACTATCAGTGC**TAG**ACGGCACCGT ATCTGCAGTGGGCACGGGGGGCCGGCCAGACAGGCAGACTGGGAGGATGGAGGACGGAGCT GCAGACGAAGGCAGGCAGCCCCAGGCACCCCAGGCAGTCTGTGTG ACATGCAGATATGCTGCCTGGACACACACTTCCAGACACACGTGCACAGGCGCAGATATGCT GCAGATATGCTGCCTGGACACACACACAGATAATGCTGCCTCAACACTCACACACGTGCAGA TATTGCCTGGACACACATGTGCACAGATATGCTGTCTGGACATGCACACACGTGCAGATA TGCTGTCCGGATACACACGCACGCACACATGCAGATATGCTGCCTGGGCACACACTTCCGGA CACACATGCACACACAGGTGCAGATATGCTGCCTGGACACGCAGACTGACGTGCTTTTGG GAGGGTGTGCCGTGAAGCCTGCAGTACGTGTGCCGTGAGGCTCATAGTTGATGAGGGACTTT CCCTGCTCCACCGTCACTCCCCCAACTCTGCCCGCCTCTGTCCCCGCCTCAGTCCCCGCCTC
CATCCCCGCCTCTGTCCCCTGGCCTTGGCGGCTATTTTTTGCCACCTGCCTTGGGTGCCCAGG
AGTCCCCTACTGCTGTGGGGTTGGGGGCACAGCAGCCCCAAGCCTGAGAGGGCTGGAG
CCCATGGCTAGTGGCTCATCCCCAGTGCATTCTCCCCCTGACACAGAGAAGGGGCCTTGGTA
TTTATATTTAAGAAATGAAGATAATATTAATAATGATGGAAGGAAGACTGGGTTGCAGGGAC

FIGURE 70

MTPSPLLLLLPPLLLGAFPPAAAARGPPKMADKVVPRQVARLGRTVRLQCPVEGDPPPLTM
WTKDGRTIHSGWSRFRVLPQGLKVKQVEREDAGVYVCKATNGFGSLSVNYTLVVLDDISPGK
ESLGPDSSSGGQEDPASQQWARPRFTQPSKMRRRVIARPVGSSVRLKCVASGHPRPDITWMK
DDQALTRPEAAEPRKKKWTLSLKNLRPEDSGKYTCRVSNRAGAINATYKVDVIQRTRSKPVL
TGTHPVNTTVDFGGTTSFQCKVRSDVKPVIQWLKRVEYGAEGRHNSTIDVGGQKFVVLPTGD
VWSRPDGSYLNKLLITRARQDDAGMYICLGANTMGYSFRSAFLTVLPDPKPPGPPVASSSSA
TSLPWPVVIGIPAGAVFILGTLLLWLCQAQKKPCTPAPAPPLPGHRPPGTARDRSGDKDLPS
LAALSAGPGVGLCEEHGSPAAPQHLLGPGPVAGPKLYPKLYTDIHTHTHTHSHTHSHVEGKV
HQHIHYQC

FIGURE 71

CCCAGCTGAGGAGCCCTGCTCAAGACACGGTCACTGGATCTGAGAAACTTCCCAGGGGACCGCATTCCAGAGTC

FIGURE 72

MVGTKAWVFSFLVLEVTSVLGRQTMLTQSVRRVQPGKKNPSIFAKPADTLESPGEWTTWFNI DYPGGKGDYERLDAIRFYYGDRVCARPLRLEARTTDWTPAGSTGOVVHGSPREGFWCLNREO RPGQNCSNYTVRFLCPPGSLRRDTERIWSPWSPWSKCSAACGQTGVQTRTRICLAEMVSLCS EASEEGQHCMGQDCTACDLTCPMGQVNADCDACMCQDFMLHGAVSLPGGAPASGAAIYLLTK TPKLLTQTDSDGRFRIPGLCPDGKSILKITKVKFAPIVLTMPKTSLKAATIKAEFVRAETPY MVMNPETKARRAGQSVSLCCKATGKPRPDKYFWYHNDTLLDPSLYKHESKLVLRKLQQHQAG EYFCKAQSDAGAVKSKVAQLIVTASDETPCNPVPESYLIRLPHDCFONATNSFYYDVGRCPV KTCAGQQDNGIRCRDAVQNCCGISKTEEREIOCSGYTLPTKVAKECSCORCTETRSIVRGRV SAADNGEPMRFGHVYMGNSRVSMTGYKGTFTLHVPODTERLVLTFVDRLOKFVNTTKVLPFN KKGSAVFHEIKMLRRKEPITLEAMETNIIPLGEVVGEDPMAELEIPSRSFYRQNGEPYIGKV KASVTFLDPRNISTATAAQTDLNFINDEGDTFPLRTYGMFSVDFRDEVTSEPLNAGKVKVHL DSTQVKMPEHISTVKLWSLNPDTGLWEEEGDFKFENORRNKREDRTFLVGNLEIRERRLFNL DVPESRRCFVKVRAYRSERFLPSEQIQGVVISVINLEPRTGFLSNPRAWGRFDSVITGPNGA CVPAFCDDQSPDAYSAYVLASLAGEELQAVESSPKFNPNAIGVPQPYLNKLNYRRTDHEDPR VKKTAFQISMAKPRPNSAEESNGPIYAFENLRACEEAPPSAAHFRFYQIEGDRYDYNTVPFN EDDPMSWTEDYLAWWPKPMEFRACYIKVKIVGPLEVNVRSRNMGGTHRRTVGKLYGIRDVRS TRDRDQPNVSAACLEFKCSGMLYDQDRVDRTLVKVIPOGSCRRASVNPMLHEYLVNHLPLAV NNDTSEYTMLAPLDPLGHNYGIYTVTDQDPRTAKEIALGRCFDGTSDGSSRIMKSNVGVALT FNCVERQVGRQSAFQYLQSTPAQSPAAGTVQGRVPSRRQQRASRGGQRQGGVVASLRFPRVA OOPLIN

FIGURE 73

CTGCAAGTTGTTAACGCCTAACACACAAGTATGTTAGGCTTCCACCAAAGTCCTCAATATACCTGAATACGCAC ACCTACCTACCCGTACGCATACATACATATGTGTATATATGTAAACTAGACAAAGATCGCAGATCATAAAGC AAGCTCTGCTTTAGTTTCCAAGAAGATTACAAAGAATTTAGAGATGTATTTGTCAAGATCCCTGTCGATTCATG $\tt CCCTTTGGGTTACGGTGTCCTCAGTGATGCAGCCCTACCCTTTGGTTTGGGGACATTATGATTTGTGTAAGACT$ CAGATTTACACGGAAGAAGGGAAAGTTTGGGATTACATGGCCTGCCAGCCGGAATCCACGGACATGACAAAATA TCTGAAAGTGAAACTCGATCCTCCGGATATTACCTGTGGAGACCCTCCTGAGACGTTCTGTGCAATGGGCAATC CCTACATGTGCAATAATGAGTGTGATGCGAGTACCCCTGAGCTGGCACACCCCCTGAGCTGATGTTTTT GAAGGAAGACATCCCTCCACATTTTGGCAGTCTGCCACTTGGAAGGAGTATCCCAAGCCTCTCCAGGTTAACAT CACTCTGTCTTGGAGCAAAACCATTGAGCTAACAGACAACATAGTTATTACCTTTGAATCTGGGCGTCCAGACC AAATGATCCTGGAGAAGTCTCTCGATTATGGACGAACATGGCAGCCCTATCAGTATTATGCCACAGACTGCTTA GATGCTTTTCACATGGATCCTAAATCCGTGAAGGATTTATCACAGCATACGGTCTTAGAAATCATTTGCACAGA AGAGTACTCAACAGGGTATACAACAAATAGCAAAATAATCCACTTTGAAATCAAAGACAGGTTCGCGCTTTTTG CTGGACCTCGCCTACGCAATATGGCTTCCCTCTACGGACAGCTGGATACAACCAAGAAACTCAGAGATTTCTTT ACAGTCACAGACCTGAGGATAAGGCTGTTAAGACCAGCCGTTGGGGAAATATTTGTAGATGAGCTACACTTGGC ACGCTACTTTTACGCGATCTCAGACATAAAGGTGCGAGGAAGGTGCAAGTGTAATCTCCATGCCACTGTATGTG TGTATGACAACAGCAAATTGACATGCGAATGTGAGCACAACACTACAGGTCCAGACTGTGGGAAATGCAAGAAG AATTATCAGGGCCGACCTTGGAGTCCAGGCTCCTATCTCCCCATCCCCAAAGGCACTGCAAATACCTGTATCCC CAGTATTCCAGTATTGGTACGAATGTCTGCGACAACGAGCTCCTGCACTGCCAGAACGGAGGGACGTGCCACA ACAACGTGCGCTGCCCGGCCGCATACACGGGCATCCTCTGCGAGAAGCTGCGGTGCGAGGGGGCTGGC AGCTGCGGCTCCGACTCTGGCCAGGGCGCCCCCGCACGCCACCCCAGCGCTGCTGCTGACCACGCTGCT GGGAACCGCCAGCCCCTGGTGTTC**TAG**GTGTCACCTCCAGCCACACCGGACGGCCTGTGCCGTGGGGAAGCA CTAAGAAGGCCTAACTGAACTAAGCCATATTTATCACCCGTGGACAGCACATCCGAGTCAAGACTGTTAATTTC TGACTCCAGAGGAGTTGGCAGCTGTTGATATTATCACTGCAAATCACATTGCCAGCTGCAGAGCATATTGTGGA GTGCATTGTGGGCATAAGGAAATCTGTTACAAGCTGCCATATTGGCCTGCTTCCGTCCCTGAATCCCTTCCAAC TGTGTAACAGCCCCCTCTAAAAGCGCAAGCCAGTCATACCCCTGTATATCTTAGCAGCACTGAGTCCAGTGCGA ATTTTTCTTGAACTACTGTAATATGTAGATTTTTTGTATTATTGCCAATTTGTGTTACCAGACAATCTGTTAAT GATTTCTCTGTAAGGGCAACGAACGTGCTGGCATCAAAGAATATCAGTTTACATATAACAAGTGTAATAAGA TTCCACCAAAGGACATTCTAAATGTTTTCTTGTTGCTTTAACACTGGAAGATTTAAAGAATAAAAACTCCTGCA ${\tt TTACTGATTTCTGTGTGGACTGAGTACATTCAGCTGACGAATTTAGTTCCCAGGAAGATGGATTGATGTTCACT}$ AAAAAA

FIGURE 74

MYLSRSLSIHALWVTVSSVMQPYPLVWGHYDLCKTQIYTEEGKVWDYMACQPESTDMTKYLK VKLDPPDITCGDPPETFCAMGNPYMCNNECDASTPELAHPPELMFDFEGRHPSTFWQSATWK EYPKPLQVNITLSWSKTIELTDNIVITFESGRPDQMILEKSLDYGRTWQPYQYYATDCLDAF HMDPKSVKDLSQHTVLEIICTEEYSTGYTTNSKIIHFEIKDRFALFAGPRLRNMASLYGQLD TTKKLRDFFTVTDLRIRLLRPAVGEIFVDELHLARYFYAISDIKVRGRCKCNLHATVCVYDN SKLTCECEHNTTGPDCGKCKKNYQGRPWSPGSYLPIPKGTANTCIPSISSIGTNVCDNELLH CQNGGTCHNNVRCLCPAAYTGILCEKLRCEEAGSCGSDSGQGAPPHGTPALLLLTTLLGTAS PLVF

FIGURE 75

CCCACGCGTCCGGTGACCTGGGCCGAGCCCTCCCGGTCGGCTAAGATTGCTGAGGAGGCGG CGGGTAGCTGGCAGGCGCCGACTTCCGAAGGCCGCCGTCCGGGCGAGGTGTCCTCATGACTT $\tt CTCTTGTGGACC{\color{blue} ATG} TCCTTTTTTTGCCTGCGTGGTACGGGTAAGGGATGGACTGCC$ CCTCTCAGCCTCTACTGATTTTTACCACACCCAAGATTTTTTGGAATGGAGGAGACGGCTCA AGAGTTTAGCCTTGCGACTGGCCCAGTATCCAGGTCGAGGGTTCTGCAGAAGGTTGTGACTTT AGTATACATTTTCTTCTTCGGGGACGTGGCCTGCATGGCTATCTGCTCCTGCCAGTGTCC AGCAGCCATGGCCTTCTGCTTCCTGGAGACCCTGTGGTGGGAATTCACAGCTTCCTATGACA CTACCTGCATTGGCCTAGCCTCCAGGCCATACGCTTTTCTTGAGTTTGACAGCATCATTCAG AAAGTGAAGTGGCATTTTAACTATGTAAGTTCCTCTCAGATGGAGTGCAGCTTGGAAAAAAT TCAGGAGGAGCTCAAGTTGCAGCCTCCAGCGGTTCTCACTCTGGAGGACACAGATGTGGCAA ATGGGGTGATGAATGGTCACACCCGATGCACTTGGAGCCTGCTCCTAATTTCCGAATGGAA CCAGTGACAGCCCTGGGTATCCTCCCTCATTCTCAACATCATGTGTGCTGCCCTGAATCT CATTCGAGGAGTTCACCTTGCAGAACATTCTTTACAGGATCCAAGGAGCTGGTTCTGCTGGT TGGACCAAACCTCG**TGA**GCCAGCCACCCTGACCCAAATGAGGAGAGCTCTGATTCTCCCAT CCGGGAGCAGTGATGTCAAACTTCTGCTGCTGGGGAAATCTCATCAGCAGGGAGCCTGTGGA AAAGGGCATGTCAGTGAAATCTGGGAATGGCTGGATTCGGAAACATCTGCCCATGTGTATTG ATGGCAGAGCTGTTGCCCACAAGCGCCTTTTATTTAGGGTAAAATTAACAAATCCATTCTAT TCCTCTGACCCATGCTTAGTACATATGACCTTTAACCCTTACATTTATATGATTCTGGGGTT GCTTCAGAAGTGTTATTTCATGAATCATTCATATGATTTGATCCCCCAGGATTCTATTTTGT TTAATGGGCTTTTCTACTAAAAGCATAAAATACTGAGGCTGATTTAGTCAGGGCAAAACCAT TTACTTTACATATTCGTTTTCAATACTTGCTGTTCATGTTACACAAGCTTCTTACGGTTTTC TTGTAACAATAATATTTTGAGTAAATAATGGGTACATTTTAACAAACTCAGTAGTACAACC TAAACTTGTATAAAAGTGTGTAAAAATGTATAGCCATTTATATCCTATGTATAAATTAAATG AAAAG

FIGURE 76

MSVIFFACVVRVRDGLPLSASTDFYHTQDFLEWRRRLKSLALRLAQYPGRGSAEGCDFSIHF SSFGDVACMAICSCQCPAAMAFCFLETLWWEFTASYDTTCIGLASRPYAFLEFDSIIQKVKW HFNYVSSSQMECSLEKIQEELKLQPPAVLTLEDTDVANGVMNGHTPMHLEPAPNFRMEPVTA LGILSLILNIMCAALNLIRGVHLAEHSLQDPRSWFCWLDQTS

FIGURE 77

TGCTTCCTGGAGACCCTGTGGTGGGAATTCACAGCTTCNTATGACACTACCTGCATTGGCNT
AGCCTCCAGGCCATACGCTTTTCTTGAGTTTGACAGCATCATTCAGAAAGTGAAGTGCATT
TTAACTATGTAAGTTCCTNTCAGATGGAGTGCAGCTTGGAAAAAATTCAGGAGGAGCTCAAG
TTGCAGCCTCCAGCGGTTCTCANTATGGAGGACACAGATGTGGCAAATGGGGT

FIGURE 78

CTCAGCGGCGCTTCCTCGTAGCGAGCCTAGTGGCGGGTGTTTGCATTGAAACGTGAGCGCGA CCCGACCTTAAAGAGTGGGGAGCAAAGGGAGGACAGAGCCCTTTAAAACGAGGCGGGTGGTG CCTGCCCCTTTAAGGGCGGGGCGTCCGGACGACTGTATCTGAGCCCCAGACTGCCCCGAGTT TCTGTCGCAGGCTGCGAGGAAAGGCCCCTAGGCTGGGTCTGGGTGCTTGGCGGCGGCGGCTT $\tt CCTCCCGGCTCGTCCTCCCGGGCCCAGAGGCACCTCGGCTTCAGTCATGCTGAGCAGAGT{\color{red} A}$ TGGAAGCACCTGACTACGAAGTGCTATCCGTGCGAGAACAGCTATTCCACGAGAGGATCCGC GAGTGTATTATCAACACTTCTGTTTGCAACACTGTACATCCTCTGCCACATCTTCCTGAC CCGCTTCAAGAAGCCTGCTGAGTTCACCACAGTGGATGATGAAGATGCCACCGTCAACAAGA TTGCGCTCGAGCTGTGCACCTTTACCCTGGCAATTGCCCTGGGTGCTGTCCTGCTCCTGCCC TTCTCCATCATCAGCAATGAGGTGCTGCTCTCCCTGCCTCGGAACTACTACATCCAGTGGCT CAACGGCTCCCTCATCCATGGCCTCTGGAACCTTGTTTTTCTCTTCCCCAACCTGTCCCTCA TCTTCCTCATGCCCTTTGCATATTTCTTCACTGAGTCTGAGGGCTTTGCTGGCTCCAGAAAG GGTGTCCTGGGCCGGGTCTATGAGACAGTGGTGATGTTGATGCTCCTCACTCTGCTGGTGCT AGGTATGGTGTGGGTGGCATCAGCCATTGTGGACAAGAACAAGGCCAACAGAGTCACTCT CTGCTCCTGGTGTGTACTCCACTGGGTCTCGCCCGCATGTTCTCCGTCACTGGGAAGCTGCT AGTCAAGCCCCGGCTGCTGGAAGACCTGGAGGAGCCGCTGTACTGCTCAGCCTTTGAGGAGG CTACACAGACAGGTCCTGGCTCTGCAGACACAGAGGGTCCTGCTGGAGAAGAGGCGGAAGGC GCCTGTCTGTGCTCATTGTGGCCATCCACATCCTGGAGCTGCTCATCGATGAGGCTGCCATG CCCCGAGGCATGCAGGGTACCTCCTTAGGCCAGGTCTCCTTCTCCAAGCTGGGCTCCTTTGG TGCCGTCATTCAGGTTGTACTCATCTTTTACCTAATGGTGTCCTCAGTTGTGGGCTTCTATA GCTCTCCACTCTTCCGGAGCCTGCGGCCCAGATGGCACGACACTGCCATGACGCAGATAATT GGGAACTGTGTCTCTCCTGGTCCTAAGCTCAGCACTTCCTGTCTTCTCTCGAACCCTGGG ACTGCAGCTGTGCGGCCAGAGCTGATCCGGGCCTTTGGGCTGGACAGACTGCCGCTGCCCGT CTCCGGTTTCCCCCAGGCATCTAGGAAGACCCAGCACCAG<mark>TGA</mark>CCTCCAGCTGGGGGTGGGA AGGAAAAACTGGACACTGCCATCTGCTGCCTAGGCCTGGAGGGAAGCCCAAGGCTACTTGG ACCTCAGGACCTGGAATCTGAGAGGGTGGCTGGCAGAGGGGAGCAGAGCCATCTGCACTATT GCATAATCTGAGCCAGAGTTTGGGACCAGGACCTCCTGCTTTTCCATACTTAACTGTGGCCT CAGCATGGGGTAGGGCTGGGTGACTGGGTCTAGCCCCTGATCCCAAATCTGTTTACACATCA ATCTGCCTCACTGCTGTTCTGGGCCATCCCCATAGCCATGTTTACATGATTTGATGTGCAAT CTTGCCTCTGGCCCAGCAGAGCCTAAGCACTGTGCTATCCTGGAGGGGCTTTGGACCACCTG AAAGACCAAGGGGATAGGGAGGAGGAGTCAGCCATCAGCAATAAAGTTGATCCCAGGGA AAAAA

FIGURE 79

MEAPDYEVLSVREQLFHERIRECIISTLLFATLYILCHIFLTRFKKPAEFTTVDDEDATVNK
IALELCTFTLAIALGAVLLLPFSIISNEVLLSLPRNYYIQWLNGSLIHGLWNLVFLFPNLSL
IFLMPFAYFFTESEGFAGSRKGVLGRVYETVVMLMLLTLLVLGMVWVASAIVDKNKANRESL
YDFWEYYLPYLYSCISFLGVLLLLVCTPLGLARMFSVTGKLLVKPRLLEDLEEQLYCSAFEE
AALTRRICNPTSCWLPLDMELLHRQVLALQTQRVLLEKRRKASAWQRNLGYPLAMLCLLVLT
GLSVLIVAIHILELLIDEAAMPRGMQGTSLGQVSFSKLGSFGAVIQVVLIFYLMVSSVVGFY
SSPLFRSLRPRWHDTAMTQIIGNCVCLLVLSSALPVFSRTLGLTRFDLLGDFGRFNWLGNFY
IVFLYNAAFAGLTTLCLVKTFTAAVRAELIRAFGLDRLPLPVSGFPQASRKTQHQ

FIGURE 80

FIGURE 81

FIGURE 82

FIGURE 83

 $\verb|MLLWVILLVLAPVSGQFARTPRPIIFLQPPWTTVFQGERVTLTCKGFRFYSPQKTKWYHRYL|\\ GKEILRETPDNILEVQESGEYRCQAQGSPLSSPVHLDFSSEMGFPHAAQANVELLGSSDLLT|\\$

FIGURE 84

CAGAAGAGGGGGCTAGCTAGCTGTCTCTGCGGACCAGGGAGACCCCCGCGCCCCCCGGTGT GAGGCGGCCTCACAGGGCCGGGTGGGCTGGCGAGCCGACGCGGCGGGGGAGGAGGCTGTGAG GAGTGTGTGGAACAGGACCCGGGACAGAGGAACCATGGCTCCGCAGAACCTGAGCACCTTTT GCCTGTTGCTGCTATACCTCATCGGGGCGGTGATTGCCGGACGAGATTTCTATAAGATCTTG GGGGTGCCTCGAAGTGCCTCTATAAAGGATATTAAAAAGGCCTATAGGAAACTAGCCCTGCA GCTTCATCCCGACCGGAACCCTGATGATCCACAAGCCCAGGAGAAATTCCAGGATCTGGGTG CTGCTTATGAGGTTCTGTCAGATAGTGAGAAACGGAAACAGTACGATACTTATGGTGAAGAA GGATTAAAAGATGGTCATCAGAGCTCCCATGGAGACATTTTTTCACACTTCTTTGGGGATTT TGGTTTCATGTTTGGAGGAACCCCTCGTCAGCAAGACAGAAATATTCCAAGAGGAAGTGATA TTATTGTAGATCTAGAAGTCACTTTGGAAGAAGTATATGCAGGAAATTTTGTGGAAGTAGTT AGAAACAAACCTGTGGCAAGGCAGGCTCCTGGCAAACGGAAGTGCAATTGTCGGCAAGAGAT GCGGACCACCCAGCTGGGCCCTGGGCGCTTCCAAATGACCCAGGAGGTGGTCTGCGACGAAT GCCCTAATGTCAAACTAGTGAATGAAGAACGAACGCTGGAAGTAGAAATAGAGCCTGGGGTG AGAGACGGCATGGAGTACCCCTTTATTGGAGAAGGTGAGCCTCACGTGGATGGGGAGCCTGG TGTACACAAATGTGACAATCTCATTAGTTGAGTCACTGGTTGGCTTTGAGATGGATATTACT CACTTGGATGGTCACAAGGTACATATTTCCCGGGATAAGATCACCAGGCCAGGAGCGAAGCT ATGGAAGAAGGGGAAGGGCTCCCCAACTTTGACAACAACAATATCAAGGGCTCTTTGATAA TCACTTTTGATGTGGATTTTCCAAAAGAACAGTTAACAGAGGAAGCGAGAGAAGGTATCAAA CAGCTACTGAAACAAGGGTCAGTGCAGAAGGTATACAATGGACTGCAAGGATAT**TGA**GAGTG TCATCATGAAATGAATAAGAGGGCTTAAGAATTTGTCCATTTGCATTCGGAAAAGAATGACC AGCAAAAGGTTTACTAATACCTCTCCCTTTGGGGATTTAATGTCTGGTGCTGCCGCCTGAGT TTCAAGAATTAAAGCTGCAAGAGGACTCCAGGAGCAAAAGAAACACAATATAGAGGGTTGGA GTTGTTAGCAATTTCATTCAAAATGCCAACTGGAGAAGTCTGTTTTTAAATACATTTTGTTG TTATTTTTA

FIGURE 85

MAPQNLSTFCLLLLYLIGAVIAGRDFYKILGVPRSASIKDIKKAYRKLALQLHPDRNPDDPQ
AQEKFQDLGAAYEVLSDSEKRKQYDTYGEEGLKDGHQSSHGDIFSHFFGDFGFMFGGTPRQQ
DRNIPRGSDIIVDLEVTLEEVYAGNFVEVVRNKPVARQAPGKRKCNCRQEMRTTQLGPGRFQ
MTQEVVCDECPNVKLVNEERTLEVEIEPGVRDGMEYPFIGEGEPHVDGEPGDLRFRIKVVKH
PIFERRGDDLYTNVTISLVESLVGFEMDITHLDGHKVHISRDKITRPGAKLWKKGEGLPNFD
NNNIKGSLIITFDVDFPKEQLTEEAREGIKQLLKQGSVQKVYNGLQGY

Important features:

Signal peptide:

amino acids 1-22

Cell attachment sequence.

amino acids 254-257

Nt-dnaJ domain signature.

amino acids 67-87

Homologous region to Nt-dnaJ domain proteins.

amino acids 26-58

N-glycosylation site.

amino acids 5-9, 261-265

Tyrosine kinase phosphorylation site.

amino acids 253-260

N-myristoylation site.

amino acids 18-24, 31-37, 93-99, 215-221

Amidation site.

amino acids 164-168

FIGURE 86

TGGGACCAGGGAACCCCGGGCCCCCCGGTGGAGNGCCTAACAGGCCGGTGGNTGCGACCGAA
GCGGCGGGCGGAGGAGGTTTTGAGGATTTTTGGAACAGGACCCGGACAGAGGAACCATGGTT
CCGCAGAACNTGAGCACNTTTTGCCTGTTGNTGNTATACTTCATCGGGGCGGTGATTGCCGG
ACGAGATTTNTATAAGATTTTGGGGTGCCTNGAAGTGCCTTNTATAAAAGGATATTAAAAAGG
CCTATAGGAAACTAGCCCTGCAGNTTTATCCCGACCGGAACCCTGATGATCCACAAGCCCAG
GAGAAATTCCAGGATTTGGGTGCTGCTTATGAGGTTNTGTCAGATAGTGAGAAACGGAAACA
GTACGATAATTATGGTGAAGAAGGATTAAAAGATGGTNATCAGAGCTCCCATGGAGACATTT
TTTCACACTTNTTTGGGGATTTTGGTTTCATGTTTGGAGGAACCCCTNGTCAGCAAGACAGA
AATATTCCAAGAG

FIGURE 87

GGCACGAGGCGGGGGGCAGTCGCGGGATGCGCCCGGGAGCCACAGCCTGAGGCCCTCAGGT CTCTGCAGGTGTCGTGGAGGAACCTAGCACCTGCCATCCTCTTCCCCAATTTGCCACTTCCA GCAGCTTTAGCCCATGAGGAGGATGTGACCGGGACTGAGTCAGGAGCCCTCTGGAAGC**ATG**G AGACTGTGGTGATTGTTGCCATAGGTGTGCTGGCCACCATCTTTCTGGCTTCGTTTGCAGCC TTGGTGCTGGTTTGCAGGCAGCGCTACTGCCGGCCGCGAGACCTGCTGCAGCGCTATGATTC TAAGCCCATTGTGGACCTCATTGGTGCCATGGAGACCCAGTCTGAGCCCTCTGAGTTAGAAC TGGACGATGTCGTTATCACCAACCCCCACATTGAGGCCCATTCTGGAGAATGAAGACTGGATC GAAGATGCCTCGGGTCTCATGTCCCACTGCATTGCCATCTTGAAGATTTGTCACACTCTGAC AGAGAAGCTTGTTGCCATGACAATGGGCTCTGGGGCCAAGATGAAGACTTCAGCCAGTGTCA GCGACATCATTGTGGTGGCCAAGCGGATCAGCCCCAGGGTGGATGATGTTGTGAAGTCGATG TACCCTCCGTTGGACCCCAAACTCCTGGACGCCCGGACGACTGCCCTGCTCTGTCAG TCACCTGGTGCTGGTGACAAGGAATGCCTGCCATCTGACGGGAGGCCTGGACTGGATTGACC AGTCTCTGTCGGCTGAGGAGCATTTGGAAGTCCTTCGAGAAGCAGCCCTAGCTTCTGAG $\tt CCAGATAAAGGCCTCCCAGGCCCTGAAGGCTTCCTGCAGGAGCAGTCTGCAATT{\color{red}{\textbf{TAG}}} TGCCT$ ACAGGCCAGCAGCATGAAGGCCCCTGCCGCCATCCCTGGATGGCTCAGCTTAGCCTT TAAAGCAGGAGATCCCCGTCAGTTTATGCCTCTTTTGCAGTTGCAAACTGTGGCTGAGT GGCAGTCTAATACTACAGTTAGGGGAGATGCCATTCACTCTCTGCAAGAGGAGTATTGAAAA $\tt CTGGTGGACTGTCAGCTTTATTTAGCTCACCTAGTGTTTTCAAGAAAATTGAGCCACCGTCT$ AAGAAATCAAGAGGTTTCACATTAAAATTAGAATTTCTGGCCTCTCTCGATCGGTCAGAATG GGTCCCTGAGGCGTCTGGGTCTCTCCTCTCCCTTGCAGGTTTGGGTTTGAAGCTGAGGAACT

FIGURE 88

METVVIVAIGVLATIFLASFAALVLVCRQRYCRPRDLLQRYDSKPIVDLIGAMETQSEPSEL ELDDVVITNPHIEAILENEDWIEDASGLMSHCIAILKICHTLTEKLVAMTMGSGAKMKTSAS VSDIIVVAKRISPRVDDVVKSMYPPLDPKLLDARTTALLLSVSHLVLVTRNACHLTGGLDWI DQSLSAAEEHLEVLREAALASEPDKGLPGPEGFLQEQSAI

FIGURE 89

GCTTCATTTCTCCCGACTCAGCTTCCCACCCTGGGCTTTCCGAGGTGCTTTCGCCGCTGTCC CCACCACTGCAGCC**ATG**ATCTCCTTAACGGACACGCAGAAAATTGGAATGGGATTAACAGGA ${\tt TTTGGAGTGTTTTTCCTGTTCTTTTGGAATGATTCTCTTTTTTGACAAAGCACTACTGGCTAT}$ TCTTCTTCCAAAAACATAAAATGAAAGCTACAGGTTTTTTTCTGGGTGGTGTATTTGTAGTC CTTATTGGTTGGCCTTTGATAGGCATGATCTTCGAAATTTATGGATTTTTTCTCTTGTTCAG GGGCTTCTTTCCTGTCGTTGTTGGCTTTATTAGAAGAGTGCCAGTCCTTGGATCCCTCCTAAAT TTACCTGGAATTAGATCATTTGTAGATAAAGTTGGAGAAAGCAACAATATGGTA**TAA**CAACA GCACAAAATTAAATTACATGAAATAGCTTGTAATGTTCTTTACAGGAGTTTAAAACGTATAG CCTACAAAGTACCAGCAGCAAATTAGCAAAGAAGCAGTGAAAACAGGCTTCTACTCAAGTGA ACTAAGAAGTCAGCAAGCAAACTGAGAGGGTGAAATCCATGTTAATGATGCTTAAGAA ACTCTTGAAGGCTATTTGTGTTTTTTCCACAATGTGCGAAACTCAGCCATCCTTAGAGAA $\tt CTGTGGTGCCTGTTTCTTTTTTTTTTTTGAAGGCTCAGGAGCATCCATAGGCATTTGCT$ TTTTAGAAGTGTCCACTGCAATGGCAAAAATATTTCCAGTTGCACTGTATCTCTGGAAGTGA TGCATGAATTCGATTGGATTGTCATTTTAAAGTATTAAAACCAAGGAAACCCCAATTTTG ATGTATGGATTACTTTTTTTTTGNGCNCAGGGCC

FIGURE 90

MISLTDTQKIGMGLTGFGVFFLFFGMILFFDKALLAIGNVLFVAGLAFVIGLERTFRFFFQK HKMKATGFFLGGVFVVLIGWPLIGMIFEIYGFFLLFRGFFPVVVGFIRRVPVLGSLLNLPGI RSFVDKVGESNNMV

Important features:

Transmembrane domains:

amino acids 12-30 (typeII), 33-52, 69-89 and 93-109

N-myristoylation sites.

amino acids 11-16, 51-56 and 116-121

Aminoacyl-transfer RNA synthetases class-II protein.

amino acids 49-59

FIGURE 91

FIGURE 92

GGCACGAGGCTGAACCCAGCCGGCTCCATCTCAGCTTCTGGTTTCTAAGTCCATGTGCCAAA $\tt CTGTGGGTAGTTATTTTTTTGAATAAGAGCGTCCACGCATC \underline{\textbf{ATG}} \texttt{GACCTCGCGGGACTGC}$ TGAAGTCTCAGTTCCTGTGCCACCTGGTCTTCTGCTACGTCTTTATTGCCTCAGGGCTAATC ATCAACACCATTCAGCTCTTCACTCTCCTCCTCTGGCCCATTAACAAGCAGCTCTTCCGGAA GATCAACTGCAGACTGTCCTATTGCATCTCAAGCCAGCTGGTGATGCTGCTGGAGTGGTGGT GCCATCGTGGTTCTCAACCACAAGTTTGAAATTGACTTTCTGTGTGGGCTGGAGCCTGTCCGA TTATCGGCTGGATGTGGTACTTCACCGAGATGGTCTTCTGTTCGCGCAAGTGGGAGCAGGAT CGCAAGACGGTTGCCACCAGTTTGCAGCACCTCCGGGACTACCCCGAGAAGTATTTTTTCCT GATTCACTGTGAGGGCACACGGTTCACGGAGAAGAAGCATGAGATCAGCATGCAGGTGGCCC GGGCCAAGGGGCTGCCTCGCCTCAAGCATCACCTGTTGCCACGAACCAAGGGCTTCGCCATC ACCGTGAGGAGCTTGAGAAATGTAGTTTCAGCTGTATATGACTGTACACTCAATTTCAGAAA TTAGGAGGATCCCACTGGAAGACATCCCTGAAGACGATGACGAGTGCTCGGCCTGGCTGCAC AAGCTCTACCAGGAGGAGGATGCCTTTCAGGAGGAGTACTACAGGACGGCCACCTTCCCAGA GACGCCCATGGTGCCCCCCGGCGCCCCTGGACCCTCGTGAACTGGCTGTTTTGGGCCTCGC TGGTGCTCTACCCTTTCTTCCAGTTCCTGGTCAGCATGATCAGGAGCGGGTCTTCCCTGACG <u>GA</u>CTCAGGGAGGTGTCACCATCCGAAGGGAACCTTGGGGGAACTGGTGGCCTCTGCATATCCT CCTTAGTGGGACACGGTGACAAAGGCTGGGTGAGCCCCTGCTGGGCACGGCGGAAGTCACGA CCTCTCCAGCCAGGGAGTCTGGTCTCAAGGCCGGATGGGGAGGAAGATGTTTTGTAATCTTT TGTGTGGTGAGTGTGAACTTTGTTCTGTGATCATAGAAAGGGTATTTTAGGCTGCAGGGGAG GGCAGGGCTGGGGACCGAAGGGGACAAGTTCCCCTTTCATCCTTTGGTGCTGAGTTTTCTGT AACCCTTGGTTGCCAGAGATAAAGTGAAAAGTGCTTTAGGTGAGATGACTAAATTATGCCTC

FIGURE 93

MDLAGLLKSQFLCHLVFCYVFIASGLIINTIQLFTLLLWPINKQLFRKINCRLSYCISSQLV
MLLEWWSGTECTIFTDPRAYLKYGKENAIVVLNHKFEIDFLCGWSLSERFGLLGGSKVLAKK
ELAYVPIIGWMWYFTEMVFCSRKWEQDRKTVATSLQHLRDYPEKYFFLIHCEGTRFTEKKHE
ISMQVARAKGLPRLKHHLLPRTKGFAITVRSLRNVVSAVYDCTLNFRNNENPTLLGVLNGKK
YHADLYVRRIPLEDIPEDDDECSAWLHKLYQEKDAFQEEYYRTGTFPETPMVPPRRPWTLVN
WLFWASLVLYPFFQFLVSMIRSGSSLTLASFILVFFVASVGVRWMIGVTEIDKGSAYGNSDS
KQKLND

FIGURE 94

 $\tt CTGAGGCGGTAGC \underline{\textbf{ATG}} GAGGGGGAGAGTACGTCGGCGGTGCTCTCGGGCTTTGTGCTCG$ GCGCACTCGCTTTCCAGCACCTCAACACGGACTCGGACACGGAAGGTTTTCTTCTTGGGGAA GTAAAAGGTGAAGCCAAGAACAGCATTACTGATTCCCAAATGGATGATGTTGAAGTTGTTTA TACAATTGACATTCAGAAATATATTCCATGCTATCAGCTTTTTAGCTTTTATAATTCTTCAG GCGAAGTAAATGAGCAAGCACTGAAGAAAATATTATCAAATGTCAAAAAGAATGTGGTAGGT AAACTTGCAGGAGCATTTTTCAAACCAAGACCTTGTTTTTCTGCTATTAACACCAAGTATAA TAACAGAAAGCTGCTCTACTCATCGACTGGAACATTCCTTATATAAACCTCAAAAAGGACTT TTTCACAGGGTACCTTTAGTGGTTGCCAATCTGGGCATGTCTGAACAACTGGGTTATAAAAC TGTATCAGGTTCCTGTATGTCCACTGGTTTTAGCCGAGCAGTACAAACACACAGCTCTAAAT TTTTTGAAGAAGATGGATCCTTAAAGGAGGTACATAAGATAAATGAAATGTATGCTTCATTA CAAGAGGAATTAAAGAGTATATGCAAAAAAGTGGAAGACAGTGAACAAGCAGTAGATAAACT AGTAAAGGATGTAAACAGATTAAAACGAGAAATTGAGAAAAGGAGAGGAGCACAGATTCAGG CAGCAAGAGAAGAACATCCAAAAAGACCCTCAGGAGAACATTTTTCTTTGTCAGGCATTA CGGACCTTTTTTCCAAATTCTGAATTTCTTCATTCATGTGTTATGTCTTTAAAAAAATAGACA TGTTTCTAAAAGTAGCTGTAACTACAACCACCATCTCGATGTAGTAGACAATCTGACCTTAA TGGTAGAACACACTGACATTCCTGAAGCTAGTCCAGCTAGTACACCACAAATCATTAAGCAT AAAGCCTTAGACTTAGATGACAGATGGCAATTCAAGAGATCTCGGTTGTTAGATACACAAGA CAAACGATCTAAAGCAAATACTGGTAGTAGTAACCAAGATAAAGCATCCAAAATGAGCAGCC CAGAAACAGATGAAGAATTGAAAAGATGAAGGGTTTTGGTGAATATTCACGGTCTCCTACA $\mathtt{TTT}\mathbf{TGA}\mathtt{TCCTTTTAACCTTACAAGGAGATTTTTTTTTTTTGGCTGATGGGTAAAGCCAAACAT$ TTCTATTGTTTTACTATGTTGAGCTACTTGCAGTAAGTTCATTTGTTTTTACTATGTTCAC CTGTTTGCAGTAATACACAGATAACTCTTAGTGCATTTACTTCACAAAGTACTTTTTCAAAC ATCAGATGCTTTTATTTCCAAACCTTTTTTTCACCTTTCACTAAGTTGTTGAGGGGAAGGCT TACACAGACACATTCTTTAGAATTGGAAAAGTGAGACCAGGCACAGTGGCTCACACCTGTAA TGGGCAACGTATTGAGACCATGTCTATTAAAAAATAAAATGGAAAAGCAAGAATAGCCTTAT TTTCAAAATATGGAAAGAAATTTATATGAAAATTTATCTGAGTCATTAAAATTCTCCTTAAG TGATACTTTTTTAGAAGTACATTATGGCTAGAGTTGCCAGATAAAATGCTGGATATCATGCA

FIGURE 95

MEGESTSAVLSGFVLGALAFQHLNTDSDTEGFLLGEVKGEAKNSITDSQMDDVEVVYTIDIQ
KYIPCYQLFSFYNSSGEVNEQALKKILSNVKKNVVGWYKFRRHSDQIMTFRERLLHKNLQEH
FSNQDLVFLLLTPSIITESCSTHRLEHSLYKPQKGLFHRVPLVVANLGMSEQLGYKTVSGSC
MSTGFSRAVQTHSSKFFEEDGSLKEVHKINEMYASLQEELKSICKKVEDSEQAVDKLVKDVN
RLKREIEKRRGAQIQAAREKNIQKDPQENIFLCQALRTFFPNSEFLHSCVMSLKNRHVSKSS
CNYNHHLDVVDNLTLMVEHTDIPEASPASTPQIIKHKALDLDDRWQFKRSRLLDTQDKRSKA
NTGSSNQDKASKMSSPETDEEIEKMKGFGEYSRSPTF

FIGURE 96

CCTTCCCCGCGTCCCCGCCTCGCCGGCCAGTCAGCTTGCCGGGTTCGCTGCCCCGCGAAACCCCGAGGTCACCA GGGGACCGTTGCCTGACGCGAGGCCCAGCTCTACTTTTCGCCCCGCGTCTCCTCCGCCTGCTCGCCTCTTCCAC CAACTCCAACTCCTCCCCCCAGCTCCACTCGCTAGTCCCCGACTCCGCCAGCCCTCGGCCCGCTGCCGTAG AAGTGCGACGTCTTTACGTGTCCAAAGGCTTCAACAAGAACGATGCCCCCCTCCACGAGATCAACGGTGATCAT TTGAAGATCTGTCCCCAGGGTTCTACCTGCTCTCAAGAGATGGAGGAGAAGTACAGCCTGCAAAGTAAAGA TGATTTCAAAAGTGTGGTCAGCGAACAGTGCAATCATTTGCAAGCTGTCTTTGCTTCACGTTACAAGAAGTTTG ATGAATTCTTCAAAGAACTACTTGAAAATGCAGAGAAATCCCTGAATGATATGTTTGTGAAGACATATGGCCAT TTATACATGCAAAATTCTGAGCTATTTAAAGATCTCTTCGTAGAGTTGAAACGTTACTACGTGGTGGGAAATGT GAACCTGGAAGAAATGCTAAATGACTTCTGGGCTCGCCTCCTGGAGCGGATGTTCCGCCTGGTGAACTCCCAGT ACCACTTTACAGATGAGTATCTGGAATGTGTGAGCAAGTATACGGAGCAGCTGAAGCCCTTCGGAGATGTCCCT CGCAAATTGAAGCTCCAGGTTACTCGTGCTTTTGTAGCAGCCCGTACTTTCGCTCAAGGCTTAGCGGTTGCGGG ${\tt AGATGTCGTGAGCAAGGTCTCCGTGGTAAACCCCACAGCCCAGTGTACCCATGCCCTGTTGAAGATGATCTACT}$ GCTCCCACTGCCGGGGTCTCGTGACTGTGAAGCCATGTTACAACTACTGCTCAAACATCATGAGAGGCTGTTTG GCCAACCAAGGGGATCTCGATTTTGAATGGAACAATTTCATAGATGCTATGCTGATGGTGGCAGAGAGGCTAGA GGGTCCTTTCAACATTGAATCGGTCATGGATCCCATCGATGTGAAGATTTCTGATGCTATTATGAACATGCAGG ATAATAGTGTTCAAGTGTCTCAGAAGGTTTTCCAGGGATGTGGACCCCCCAAGCCCCTCCCAGCTGGACGAATT TCTCGTTCCATCTCTGAAAGTGCCTTCAGTGCTCGCTTCAGACCACATCACCCCGAGGAACGCCCAACCACAGC AGCTGGCACTAGTTTGGACCGACTGGTTACTGATGTCAAGGAGAAACTGAAACAGGCCAAGAAATTCTGGTCCT CCCTTCCGAGCAACGTTTGCAACGATGAGAGGATGGCTGCAGGAAACGGCAATGAGGATGACTGTTGGAATGGG AAAGGCAAAAGCAGGTACCTGTTTGCAGTGACAGGAAATGGATTAGCCAACCAGGGCAACAACCCAGAGGTCCA GGTTGACACCAGCAAACCAGACATACTGATCCTTCGTCAAATCATGGCTCTTCGAGTGATGACCAGCAAGATGA AGAATGCATACAATGGGAACGACGTGGACTTCTTTGATATCAGTGATGAAAGTAGT 3GAGAAGGAAGTGGAAGT GGCTGTGAGTATCAGCAGTGCCCTTCAGAGTTTGACTACAATGCCACTGACCATGCTGGGAAGAGTGCCAATGA ${\tt GAAAGCCGACAGTGCTGGTGTCCTGGGGCACAGGCCTACCTCCTCACTGTTCTTGCATCTTGTTCCTGG}$ ${\tt TTATGCAGAGAGAGTGGAGA} \underline{{\tt TAA}} {\tt TTCTCAAACTCTGAGAAAAAGTGTTCATCAAAAAGTTAAAAGGCACCAGTT}$ ${\tt CGCCTTGTTTCTTACAAGCAAACCAGGGTCCCTTCTTGGCACGTAACATGTACGTATTTCTGAAATATTAAATA}$ GCTGTACAGAAGCAGGTTTTATTTATCATGTTATCTTATTAAAAGAAAAAGCCCAAAAAGC

FIGURE 97

MARFGLPALLCTLAVLSAALLAAELKSKSCSEVRRLYVSKGFNKNDAPLHEINGDHLKICPQ
GSTCCSQEMEEKYSLQSKDDFKSVVSEQCNHLQAVFASRYKKFDEFFKELLENAEKSLNDMF
VKTYGHLYMQNSELFKDLFVELKRYYVVGNVNLEEMLNDFWARLLERMFRLVNSQYHFTDEY
LECVSKYTEQLKPFGDVPRKLKLQVTRAFVAARTFAQGLAVAGDVVSKVSVVNPTAQCTHAL
LKMIYCSHCRGLVTVKPCYNYCSNIMRGCLANQGDLDFEWNNFIDAMLMVAERLEGPFNIES
VMDPIDVKISDAIMNMQDNSVQVSQKVFQGCGPPKPLPAGRISRSISESAFSARFRPHHPEE
RPTTAAGTSLDRLVTDVKEKLKQAKKFWSSLPSNVCNDERMAAGNGNEDDCWNGKGKSRYLF
AVTGNGLANQGNNPEVQVDTSKPDILILRQIMALRVMTSKMKNAYNGNDVDFFDISDESSGE
GSGSGCEYQQCPSEFDYNATDHAGKSANEKADSAGVRPGAQAYLLTVFCILFLVMQREWR

FIGURE 98

FIGURE 99

MKVLISSLLLLLPLMLMSMVSSSLNPGVARGHRDRGQASRRWLQEGGQECECKDWFLRAPRR KFMTVSGLPKKQCPCDHFKGNVKKTRHQRHHRKPNKHSRACQQFLKQCQLRSFALPL

FIGURE 100

FIGURE 101

MAVLVLRLTVVLGLLVLFLTCYADDKPDKPDDKPDDSGKDPKPDFPKFLSLLGTEIIENAVE FILRSMSRSTGFMEFDDNEGKHSSK

FIGURE 102

GGACGCCAGCGCCTGCAGAGGCTGAGCAGGGAAAAAGCCAGTGCCCCAGCGGAAGCACAGCT CAGAGCTGGTCTGCCATGGACATCCTGGTCCCACTCCTGCAGCTGCTGGTGCTGCTTCTTAC CCTGCCCCTGCACCTCATGGCTCTGCTGGGCTGCTGGCAGCCCCTGTGCAAAAGCTACTTCCCCTACCTGATGGCCGTGCTGACTCCCAAGAGCCAACCGCAAGATGGAGAGCAAGAAACGGGAG CTCTTCAGCCAGATAAAGGGGCTTACAGGAGCCTCCGGGAAAGTGGCCCTACTGGAGCTGGG CTGCGGAACCGGAGCCAACTTTCAGTTCTACCCACCGGGCTGCAGGGTCACCTGCCTAGACC CAAATCCCCACTTTGAGAAGTTCCTGACAAAGAGCATGGCTGAGAACAGGCACCTCCAATAT GGAGAGTACTGAGACCGGGAGGTGTGCTCTTTTTCTGGGAGCATGTGGCAGAACCATATGGA AGCTGGGCCTTCATGTGGCAGCAAGTTTTCGAGCCCACCTGGAAACACATTGGGGATGGCTG CTGCCTCACCAGAGAGCCTGGAAGGATCTTGAGAACGCCCAGTTCTCCGAAATCCAAATGG AACGACAGCCCCCTCCCTTGAAGTGGCTACCTGTTGGGCCCCACATCATGGGAAAGGCTGTC AAACAATCTTTCCCAAGCTCCAAGGCACTCATTTGCTCCTTCCCCAGCCTCCAATTAGAACA AGCCACCACCAGCCTATCTATCTTCCACTGAGAGGGACCTAGCAGAATGAGAGAAGACATT CATGTACCACCTACTAGTCCCTCTCCCCAACCTCTGCCAGGGCAATCTCTAACTTCAATC CCGCCTTCGACAGTGAAAAAGCTCTACTTCTACGCTGACCCAGGGAGGAAACACTAGGACCC TGTTGTATCCTCAACTGCAAGTTTCTGGACTAGTCTCCCAACGTTTGCCTCCCAATGTTGTC CCTTTCCTTCGTTCCCATGGTAAAGCTCCTCTCGCTTTCCTCCTGAGGCTACACCCATGCGT CTCTAGGAACTGGTCACAAAAGTCATGGTGCCTGCATCCCTGCCAAGCCCCCCTGACCCTCT CTCCCCACTACCACCTTCTTCCTGAGCTGGGGGCACCAGGGAGAATCAGAGATGCTGGGGAT AACCACG

FIGURE 103

MDILVPLLQLLVLLLTLPLHLMALLGCWQPLCKSYFPYLMAVLTPKSNRKMESKKRELFSQI KGLTGASGKVALLELGCGTGANFQFYPPGCRVTCLDPNPHFEKFLTKSMAENRHLQYERFVV APGEDMRQLADGSMDVVVCTLVLCSVQSPRKVLQEVRRVLRPGGVLFFWEHVAEPYGSWAFM WQQVFEPTWKHIGDGCCLTRETWKDLENAQFSEIQMERQPPPLKWLPVGPHIMGKAVKQSFP SSKALICSFPSLQLEQATHQPIYLPLRGT

FIGURE 104

GTGGGATTTATTTGAGTGCAAGATCGTTTTCTCAGTGGTGGTAGAAGTTGCCTCATCGCAGG CAGATGTTGGGGCTTTGTCCGAACAGCTCCCCTCTGCCAGCTTCTGTAGATAAGGGTTAAAA ACTAATATTTATATGACAGAAGAAAAG**ATG**TCATTCCGTAAAGTAAACATCATCTTGG TCCTGGCTGTTGCTCTCTTACTGGTTTTGCACCATAACTTCCTCAGCTTGAGCAGTTTG TTAAGGAATGAGGTTACAGATTCAGGAATTGTAGGGCCTCAACCTATAGACTTTGTCCCAAA TGCTCTCCGACATGCAGTAGATGGGAGACAAGAGGGAGATTCCTGTGGTCATCGCTGCATCTG AAGACAGGCTTGGGGGGGCCATTGCAGCTATAAACAGCATTCAGCACAACACTCGCTCCAAT GTGATTTTCTACATTGTTACTCTCAACAATACAGCAGACCATCTCCGGTCCTGGCTCAACAG AAGTAAAGGAGGATCCTGACCAGGGGGAATCCATGAAACCTTTAACCTTTGCAAGGTTCTAC TTGCCAATTCTGGTTCCCAGCGCAAAGAAGGCCATATACATGGATGATGATGTAATTGTGCA AGGTGATATTCTTGCCCTTTACAATACAGCACTGAAGCCAGGACATGCAGCTGCATTTTCAG AAGATTGTGATTCAGCCTCTACTAAAGTTGTCATCCGTGGAGCAGGAAACCAGTACAATTAC ATTGGCTATCTTGACTATAAAAAGGAAAGAATTCGTAAGCTTTCCATGAAAGCCAGCACTTG CTCATTTAATCCTGGAGTTTTTGTTGCAAACCTGACGGAATGGAAACGACAGAATATAACTA ACCAACTGGAAAAATGGATGAAACTCAATGTAGAAGAGGGACTGTATAGCAGAACCCTGGCT GGTAGCATCACAACACCTCCTCTGCTTATCGTATTTTATCAACAGCACTCTACCATCGATCC TATGTGGAATGTCCGCCACCTTGGTTCCAGTGCTGGAAAACGATATTCACCTCAGTTTGTAA AGGCTGCCAAGTTACTCCATTGGAATGGACATTTGAAGCCATGGGGAAGGACTGCTTCATAT ACTGATGTTTGGGAAAAATGGTATATTCCAGACCCAACAGGCAAATTCAACCTAATCCGAAG ATATACCGAGATCTCAAACATAAAG**TGA**AACAGAATTTGAACTGTAAGCAAGCATTTCTCAG GAAGTCCTGGAAGATAGCATGCATGGGAAGTAACAGTTGCTAGGCTTCAATGCCTATCGGTA GCAAGCCATGGAAAAAGATGTGTCAGCTAGGTAAAGATGACAAACTGCCCTGTCTGGCAGTC AGCTTCCCAGACAGACTATAGACTATAAATATGTCTCCATCTGCCTTACCAAGTGTTTTCTT ACATTTTTC

FIGURE 105

MSFRKVNIIILVLAVALFLLVLHHNFLSLSSLLRNEVTDSGIVGPQPIDFVPNALRHAVDGR QEEIPVVIAASEDRLGGAIAAINSIQHNTRSNVIFYIVTLNNTADHLRSWLNSDSLKSIRYK IVNFDPKLLEGKVKEDPDQGESMKPLTFARFYLPILVPSAKKAIYMDDDVIVQGDILALYNT ALKPGHAAAFSEDCDSASTKVVIRGAGNQYNYIGYLDYKKERIRKLSMKASTCSFNPGVFVA NLTEWKRQNITNQLEKWMKLNVEEGLYSRTLAGSITTPPLLIVFYQQHSTIDPMWNVRHLGS SAGKRYSPQFVKAAKLLHWNGHLKPWGRTASYTDVWEKWYIPDPTGKFNLIRRYTEISNIK

FIGURE 106

FIGURE 107

CGACGCTCTAGCGGTTACCGCTGCGGGCTGGCTGGGCTAGTGGGGCTGCGCGCTGCCACG TCCGCATCTCCATCGCCTGCAGTAAGGGCGGCCGCGGGGAGCCTTTGAGGGGAACGACT TGTCGGAGCCCTAACCAGGGGTGTCTCTGAGCCTGGTGGGATCCCCGGAGCGTCACATCACT TTCCGATCACTTCAAAGTGGTTAAAAACTAATATTTATATGACAGAAGAAAAAGATGTCATT CCGTAAAGTAAACATCATCTTGGTCCTGGGCTGTTGCTCTTCTTACTGGTTTTGCAC CATAACTTCCTCAGCTTGAGGCAGTTTGTTAAGGAATGAGGTTACAGATTCAGGAATTGTAG GGCCTCAACCTATAGGACTTTGTCCCAAATGCTCTCCGACATGCAGTAGATGGGAGACAAGA ACAGCATTCAGCACAACACTCGCTCCAATGTGATTTTCTACATTGTTACTCTCAACAATACA GCAGACCATCTCCGGTCCTGGGCTCAACAGTGATTCCCTGAAAAGCATCAGATACAAAATTG TCAATTTTGACCCTAAACTTTTGGAAGGAAAAGTAAAGGAGGATCCTGACCAGGGGGAATCC ATGAAACCTTTAACCTTTGCAAGGTTCTACTTGCCAATTCTGGGTTCCCAGCGCAAAGAAGG CCATATACATGGATGATGTAATTGTGCAAGGTGATATTCTTGCCCTTTACAATACAGCA CTGAAGCCAGGACATGCAGCTGCATTTTCAGAAGATTGTGATTCAGCCTCTACTAAAGTTGT CATCCGTGGAGCAGGAAACCAGTACAATTACATTGGCTATCTTGACTATAAAAAGGAAAGAA TTCGTAAGCTTTCCATGAAAGCCAGCACTTGCTCATTTAATCCTGGAGTTTTTGTTGCAAAC CTGACGGAATGGAAACGACAGAATATAACTAACCAACTGGAAAAATGGATGAAACTCAATGT AGAAGAGGGACTGTATAGCAGAACCCTGGCTGGTAGCATCACAACACCTCCTCTGCTTATCG TATTTTATCAACAGCACTCTACCATCGATCCTATGTGGAATGTCCGCCACCTTGGTTCCAGT GCTGGAAAACGATATTCACCTCAGTTTGTAAAGGCTGCCAAGTTACTCCATTGGAATGGACA TTTGAAGCCATGGGGAAGGACTGCTTCATATACTGATGTTTGGGGAAAAATGGTATATTCCA GACCCAACAGGCAAATTCAACCTAATCCGAAGATATACCGAGATCTCAAACATAAAGTGAAA CAGAATTTGAACTGTAAGCAAGCATTTCTCAGGAAGTCCTGGAAGATAGCATGCGTGGGAAG TAACAGTTGCTAGGCTTCAATGCCTATCGGTAGCAAGCCATGGAAAAAGATGTGTCAGCTAG ATGTCTCCATCTGCCTTACCAAGTGTTTTCTTACTACAATGCTGAATGACTGGAAAGAAGAA CTGATATGGCTAGTTCAGCTAGCTGGTACAGATAATTCAAAACTGCTGTTGGTTTTAATTTT AAAAA

FIGURE 108

FIGURE 109

MGAAISQGALIAIVCNGLVGFLLLLLWVILCWACHSRLPTLTLSLNPVPTPALAPVLRRPHH PRSPAMKAATCCSPEGPWPSLEPRT

FIGURE 110

GTTTGAATTCCTTCAACTATACCCACAGTCCAAAAGCAGACTCACTGTGTCCCAGGCTACCA GTTCCTCCAAGCAAGTCATTTCCCTTATTTAACCGATGTGTCCCTCAAACACCTGAGTGCTA CTCCCTATTTGCATCTGTTTTGATAAATGATGTTGACACCCTCCACCGAATTCTAAGTGGAA $\verb"TC" \underline{\textbf{ATG}} \texttt{TCGGGGAAGAGATACAATCCTTGGCCTGTGTATCCTCGCATTAGCCTTGTCTTTGGCC"$ ATGATGTTTACCTTCAGATTCATCACCACCCTTCTGGTTCACATTTTCATTTCATTGGTTAT TTTGGGATTGTTTGTCTGCGGTGTTTTATGGTGGCTGTATTATGACTATACCAACGACC TCCACAGGCATCACGGCAGTGCTGCTCTTGATTTTTGTTCTCAGAAAGAGAATAAAATT GACAGTTGAGCTTTTCCAAATCACAAATAAAGCCATCAGCAGTGCTCCCTTCCTGCTGTTCC AGCCACTGTGGACATTTGCCATCCTCATTTTCTTCTGGGTCCTCTGGGTGGCTGCTGCTG AGCCTGGGAACTGCAGGAGCTGCCCAGGTTATGGAAGGCGGCCAAGTGGAATATAAGCCCCT TTCGGGCATTCGGTACATGTGGTCGTACCATTTAATTGGCCTCATCTGGACTAGTGAATTCA TCCTTGCGTGCCAGCAAATGACTATAGCTGGGGCAGTGGTTACTTGTTATTTCAACAGAAGT AAAAATGATCCTCCTGATCATCCCATCCTTTCGTCTCTCCCATTCTCTTCTTCTACCATCA ${\tt AGGAACCGTTGTGAAAGGGTCATTTTTAATCTCTGTGGTGAGGATTCCGAGAATCATTGTCA}$ TGTACATGCAAAACGCACTGAAAGAACAGCAGCATGGTGCATTGTCCAGGTACCTGTTCCGA TGCTGCTACTGCTGTTTCTGGTGTCTTGACAAATACCTGCTCCATCTCAACCAGAATGCATA TACTACAACTGCTATTAATGGGACAGATTTCTGTACATCAGCAAAAGATGCATTCAAAATCT TGTCCAAGAACTCAAGTCACTTTACATCTATTAACTGCTTTGGAGACTTCATAATTTTTCTA GGAAAGGTGTTAGTGGTGTTTTCACTGTTTTTTGGAGGACTCATGGCTTTTAACTACAATCG ATAGTTTTTTATCTGTGTTTGAAACTGTGCTGGATGCACTTTTCCTGTGTTTTGCTGTTGAT CTGGAAACAAATGATGGATCGTCAGAAAAGCCCTACTTTATGGATCAAGAATTTCTGAGTTT CGTAAAAAGGAGCAACAAATTAAACAATGCAAGGGCACAGCAGGACAAGCACTCATTAAGGA ${\tt ATGAGGAGGGAACAGAACTCCAGGCCATTGTGAGA} \underline{{\tt TAG}} {\tt ATACCCATTTAGGTATCTGTACCT}$ GGAAAACATTTCCTTCTAAGAGCCATTTACAGAATAGAAGATGAGACCACTAGAGAAAAGTT AGTGAATTTTTTTTTAAAAGACCTAATAAACCCTATTCTTCCTCAAAA

FIGURE 111

MSGRDTILGLCILALALSLAMMFTFRFITTLLVHIFISLVILGLLFVCGVLWWLYYDYTNDL
SIELDTERENMKCVLGFAIVSTGITAVLLVLIFVLRKRIKLTVELFQITNKAISSAPFLLFQ
PLWTFAILIFFWVLWVAVLLSLGTAGAAQVMEGGQVEYKPLSGIRYMWSYHLIGLIWTSEFI
LACQQMTIAGAVVTCYFNRSKNDPPDHPILSSLSILFFYHQGTVVKGSFLISVVRIPRIIVM
YMQNALKEQQHGALSRYLFRCCYCCFWCLDKYLLHLNQNAYTTTAINGTDFCTSAKDAFKIL
SKNSSHFTSINCFGDFIIFLGKVLVVCFTVFGGLMAFNYNRAFQVWAVPLLLVAFFAYLVAH
SFLSVFETVLDALFLCFAVDLETNDGSSEKPYFMDQEFLSFVKRSNKLNNARAQQDKHSLRN
EEGTELQAIVR

FIGURE 112

GTTCGATTAGCTCCTCTGAGAAGAAGAGAAAAGGTTCTTGGACCTCTCCCTGTTTCCTT TGTGGTGAAAATTTTTTGAAAAAAAATTGCCTTCTTCAAACAAGGGTGTCATTCTGATATT $\mathtt{T}\underline{\mathbf{ATG}}\mathtt{AGGACTGTTGTTCTCACTATGAAGGCATCTGTTATTGAAATGTTCCTTGTTTTGCTGG}$ TGACTGGAGTACATTCAAACAAAGAAACGGCAAAGAAGATTAAAAGGCCCAAGTTCACTGTG CCTCAGATCAACTGCGATGTCAAAGCCGGAAAGATCATCGATCCTGAGTTCATTGTGAAATG TCCAGCAGGATGCCAAGACCCCAAATACCATGTTTATGGCACTGACGTGTATGCATCCTACT CCAGTGTGTGTGGCGCTGCCGTACACAGTGGTGTGCTTGATAATTCAGGAGGGAAAATACTT GCTTCTGGCTGTCACTGTAGCTGTGGCCACCCCCACCACCTTGCCAAGGCCATCCCCTTCTG CTGCTTCTACCACCAGCATCCCCAGACCACAATCAGTGGGCCACAGGAGCCAGGAGATGGAT CTCTGGTCCACTGCCACCTACACAAGCAGCCAAAACAGGCCCAGAGCTGATCCAGGTATCCA AAGGCAAGATCCTTCAGGAGCTGCCTTCCAGAAACCTGTTGGAGCGGATGTCAGCCTGGGAC TTGTTCCAAAAGAAGAATTGAGCACACAGTCTTTGGAGCCAGTATCCCTGGGAGATCCAAAC TGCAAAATTGACTTGTCGTTTTTAATTGATGGGAGCACCAGCATTGGCAAACGGCGATTCCG GGGCTCCCAATGTGGTGGTGATGGTGGATGGCCCACGGACAAAGTGGAGGAGGCT TCAAGACTTGCGAGAGAGTCAGGAATCAACATTTTCTTCATCACCATTGAAGGTGCTGCTGA AAATGAGAAGCAGTATGTGGTGGAGCCCAACTTTGCAAACAAGGCCGTGTGCAGAACAAACG TTGTGACCAACCTCACCAAAGAGTTTGAGATTTCCGACACGGGACACGCGCATCGGGGCCGTG CTCACCCCGCCAGAGACCACTCCTTCTTTGTGGACGAGTTTGACAACCTCCATCAGTATGTC CAGGCAGAGCACCAGCAAGTGCTGCTTTACTAACTGACGTGTTGGACCACCCCACCGCTTAA TGGGGCACGCACGGTGCATCAAGTCTTGGGCAGGGCATGGAGAAACAAATGTCTTGTTATTA TTCTTTGCCATCATGCTTTTTCATATTCCAAAACTTGGAGTTACAAAGATGATCACAAACGT

FIGURE 113

MRTVVLTMKASVIEMFLVLLVTGVHSNKETAKKIKRPKFTVPQINCDVKAGKIIDPEFIVKC
PAGCQDPKYHVYGTDVYASYSSVCGAAVHSGVLDNSGGKILVRKVAGQSGYKGSYSNGVQSL
SLPRWRESFIVLESKPKKGVTYPSALTYSSSKSPAAQAGETTKAYQRPPIPGTTAQPVTLMQ
LLAVTVAVATPTTLPRPSPSAASTTSIPRPQSVGHRSQEMDLWSTATYTSSQNRPRADPGIQ
RQDPSGAAFQKPVGADVSLGLVPKEELSTQSLEPVSLGDPNCKIDLSFLIDGSTSIGKRRFR
IQKQLLADVAQALDIGPAGPLMGVVQYGDNPATHFNLKTHTNSRDLKTAIEKITQRGGLSNV
GRAISFVTKNFFSKANGNRSGAPNVVVVMVDGWPTDKVEEASRLARESGINIFFITIEGAAE
NEKQYVVEPNFANKAVCRTNGFYSLHVQSWFGLHKTLQPLVKRVCDTDRLACSKTCLNSADI
GFVIDGSSSVGTGNFRTVLQFVTNLTKEFEISDTDTRIGAVQYTYEQRLEFGFDKYSSKPDI
LNAIKRVGYWSGGTSTGAAINFALEQLFKKSKPNKRKLMILITDGRSYDDVRIPAMAAHLKG
VITYAIGVAWAAQEELEVIATHPARDHSFFVDEFDNLHQYVPRIIQNICTEFNSQPRN

FIGURE 114

 ${\tt CAGGATGAACTGGTTGCAGTGGCTGCTGCTGCGGGGGCGCTGAGAGGACACGAGCTCT} \underline{{\tt A}}$ $\underline{\textbf{TG}} \texttt{CCTTTCCGGCTGCTCATCCCGCTCGGCCTCCTGTGCGCGCTGCTGCCTCAGCACCATGGT}$ GCGCCAGGTCCCGACGGCTCCGCCCAGATCCCGCCCACTACAGTTTTTCTCTGACTCTAAT TGATGCACTGGACACCTTGCTGATTTTGGGGGAATGTCTCAGAATTCCAAAGAGTGGTTGAAG CGAGTGGTAGGAGGACTCCTGTCTGCTCATCTGCTCTCCAAGAAGGCTGGGGTGGAAGTAGA GGCTGGATGGCCCTGTTCCGGGCCTCTCCTGAGAATGGCTGAGGAGGCGGCCCGAAAACTCC AACCCAGGAGAGACCCCTGTCACCTGTACGGCAGGGATTGGGACCTTCATTGTTGAATTTGC CACCCTGAGCAGCCTCACTGGTGACCCGGTGTTCGAAGATGTGGCCAGAGTGGCTTTGATGC GCCTCTGGGAGAGCCGGTCAGATATCGGGCTGGTCGGCAACCACATTGATGTGCTCACTGGC AAGTGGGTGGCCCAGGACGCAGGCATCGGGGCTGGCGTGGACTCCTACTTTGAGTACTTGGT GAAAGGAGCCATCCTGCTTCAGGATAAGAAGCTCATGGCCATGTTCCTAGAGTATAACAAAG CCATCCGGAACTACACCCGCTTCGATGACTGGTACCTGTGGGTTCAGATGTACAAGGGGACT GTGTCCATGCCAGTCTTCCAGTCCTTGGAGGCCTACTGGCCTGGTCTTCAGAGCCTCATTGG AGACATTGACAATGCCATGAGGACCTTCCTCAACTACTACACTGTATGGAAGCAGTTTGGGG GGCTCCCGGAATTCTACAACATTCCTCAGGGATACACAGTGGAGAAGCGAGAGGGCTACCCA CTTCGGCCAGAACTTATTGAAAGCGCAATGTACCTCTACCGTGCCACGGGGGATCCCACCCT CCTAGAACTCGGAAGAGGTGCTGTGGAATCCATTGAAAAAATCAGCAAGGTGGAGTGCGGAT TTGCAACAATCAAAGATCTGCGAGACCACAAGCTGGACAACCGCATGGAGTCGTTCTTCCTG GTCCACCTTCGACGCGGTGATCACCCCCTATGGGGGGTGCATCCTGGGGGGGTGCA TCTTCAACACAGAAGCTCACCCCATCGACCTTGCCGCCCTGCACTGCTGCCAGAGGCTGAAG GAAGAGCAGTGGGAGGAGCTTGATGAGGGAATTCTACTCTCAAACGGAGCAGGTC GAAATTTCAGAAAAACACTGTTAGTTCGGGGCCATGGGAACCTCCAGCAAGGCCAGGAACAC $\tt CTTCTCAGCTGCCCCAGTCAGCCCTTCACCTCCAAGTTGGCATTACTGGGACAGGTTTTCCT$ AATTGCTTTTGGCTATCATAAAA

FIGURE 115

MPFRLLIPLGLLCALLPQHHGAPGPDGSAPDPAHYSFSLTLIDALDTLLILGNVSEFQRVVE
VLQDSVDFDIDVNASVFETNIRVVGGLLSAHLLSKKAGVEVEAGWPCSGPLLRMAEEAARKL
LPAFQTPTGMPYGTVNLLHGVNPGETPVTCTAGIGTFIVEFATLSSLTGDPVFEDVARVALM
RLWESRSDIGLVGNHIDVLTGKWVAQDAGIGAGVDSYFEYLVKGAILLQDKKLMAMFLEYNK
AIRNYTRFDDWYLWVQMYKGTVSMPVFQSLEAYWPGLQSLIGDIDNAMRTFLNYYTVWKQFG
GLPEFYNIPQGYTVEKREGYPLRPELIESAMYLYRATGDPTLLELGRDAVESIEKISKVECG
FATIKDLRDHKLDNRMESFFLAETVKYLYLLFDPTNFIHNNGSTFDAVITPYGECILGAGGY
IFNTEAHPIDLAALHCCQRLKEEQWEVEDLMREFYSLKRSRSKFQKNTVSSGPWEPPARPGT
LFSPENHDQARERKPAKQKVPLLSCPSQPFTSKLALLGQVFLDSS

FIGURE 116

AAAGTTACATTTTCTCTGGAACTCTCCTAGGCCACTCCCTGCTGATGCAACATCTGGGTTTG GGCAGAAAGGAGGGTGCTTCGGAGCCCGCCCTTTCTGAGCTTCCTGGGCCGGCTCTAGAACA GAGATGGACAGAATGCTTTATTTTGGAAAGAAACAATGTTCTAGGTCAAACTGAGTCTACCA $\textbf{A\textbf{ATG}} \textbf{CAGACTTTCACAATGGTTCTAGAAGAAATCTGGACAAGTCTTTTCATGTGGTTTTTCT}$ TCTGTACTCTCAACCAACATGAAGCATCTCTTGATGTGGAGCCCAGTGATCGCGCCTGGAGA AACAGTGTACTATTCTGTCGAATACCAGGGGGAGTACGAGGCCTGTACACGAGCCACATCT GGATCCCCAGCAGCTGGTGCTCACTGAAGGTCCTGAGTGTGATGTCACTGATGACATC ACGGCCACTGTGCCATACAACCTTCGTGTCAGGGCCACATTGGGCTCACAGACCTCAGCCTG GAGCATCCTGAAGCATCCCTTTAATAGAAACTCAACCATCCTTACCCGACCTGGGATGGAGA TCACCAAAGATGGCTTCCACCTGGTTATTGAGCTGGAGGACCTGGGGCCCCAGTTTGAGTTC CTTGTGGCCTACTGGAGGAGGGAGCCTGGTGCCGAGGAACATGTCAAAATGGTGAGGAGTGG GGGTATTCCAGTGCACCTAGAAACCATGGAGCCAGGGGCTGCATACTGTGTGAAGGCCCAGA GGAGAGGCCATTCCCCTGGTACTGGCCCTGTTTGCCTTTGTTGGCTTCATGCTGATCCTTGT GGTCGTGCCACTGTTCGTCTGGAAAATGGGCCGGCTGCTCCAGTACTCCTGTTGCCCCGTGG TGGTCCTCCCAGACACCTTGAAAATAACCAATTCACCCCAGAAGTTAATCAGCTGCAGAAGG GAGGAGGTGGATGCCTGTGCCACGGCTGTGATGTCTCCTGAGGAACTCCTCAGGGCCTGGAT $\tt CTCA \underline{TAG} \tt GTTTGCGGAAGGGCCCAGGTGAAGCCGAGAACCTGGTCTGCATGACATGGAAACC$ ATGAGGGGACAAGTTGTGTTTCTGTTTTCCGCCACGGACAAGGGATGAGAAGTAGGAAGA GCCTGTTGTCTACAAGTCTAGAAGCAACCATCAGAGGCAGGGTGGTTTGTCTAACAGAACAC CTGGGAAAAGTGACTTCATCCCTTCGGTCCTAAGTTTTCTCATCTGTAATGGGGGAATTACC TGTTTCTGGAGAGCAGGACATAAATGTATGATGAGAATGATCAAGGACTCTACACACTGGGT GGCTTGGAGAGCCCACTTTCCCAGAATAATCCTTGAGAGAAAAGGAATCATGGGAGCAATGG TGTTGAGTTCACTTCAAGCCCAATGCCGGTGCAGAGGGGAATGGCTTAGCGAGCTCTACAGT AGGTGACCTGGAGGAAGGTCACAGCCACACTGAAAATGGGATGTGCATGAACACGGAGGATC TGTTGGTAAAGTACAGAATTCAGCAAATAAAAAGGGCCACCCTGGCCAAAAGCGGTAAAAAA AAAAAAAA

FIGURE 117

MQTFTMVLEEIWTSLFMWFFYALIPCLLTDEVAILPAPQNLSVLSTNMKHLLMWSPVIAPGE TVYYSVEYQGEYESLYTSHIWIPSSWCSLTEGPECDVTDDITATVPYNLRVRATLGSQTSAW SILKHPFNRNSTILTRPGMEITKDGFHLVIELEDLGPQFEFLVAYWRREPGAEEHVKMVRSG GIPVHLETMEPGAAYCVKAQTFVKAIGRYSAFSQTECVEVQGEAIPLVLALFAFVGFMLILV VVPLFVWKMGRLLQYSCCPVVVLPDTLKITNSPQKLISCRREEVDACATAVMSPEELLRAWIS

Important features:

Signal peptide:

amino acids 1-29

Transmembrane domain:

amino acids 230-255

N-glycosylation sites.

amino acids 40-43 and 134-137

Tissue factor proteins homology.

amino acids 92-119

Integrins alpha chain protein homology.

amino acids 232-262

FIGURE 118

FIGURE 119

 ${\tt CGGACGCGTGGGCCACCTCCGGAACAAGCC} \underline{\textbf{ATG}} {\tt GTGGCGGCGACGGTGGCAGCGGCGTG}$ GCTGCTCCTGTGGGCTGCGCCTGCGCGCAGCAGGAGCAGGACTTCTACGACTTCAAGGCGG TCAACATCCGGGGCAAACTGGTGTCGCTGGAGAAGTACCGCGGATCGGTGTCCCTGGTGGTG AATGTGGCCAGCGAGTGCGGCTTCACAGACCAGCACTACCGAGCCCTGCAGCAGCTGCAGCG AGACCTGGGCCCCCACCACTTTAACGTGCTCGCCTTCCCCTGCAACCAGTTTGGCCAACAGG AGCCTGACAGCAACAAGGAGATTGAGAGCTTTGCCCGCCGCCCCCACCTACAGTGTCTCATTCCCC ATGTTTAGCAAGATTGCAGTCACCGGTACTGGTGCCCATCCTGCCTTCAAGTACCTGGCCCA GACTTCTGGGAAGGAGCCCACCTGGAACTTCTGGAAGTACCTAGTAGCCCCAGATGGAAAGG TGGTAGGGGCTTGGGACCCAACTGTGTCAGTGGAGGGGGTCAGACCCCAGATCACAGCGCTC GTGAGGAAGCTCATCCTACTGAAGCGAGAAGACTTA**TAA**CCACCGCGTCTCCTCCTCCACCA CCTCATCCCGCCCACCTGTGTGGGGGCTGACCAATGCAAACTCAAATGGTGCTTCAAAGGGAG AGACCCACTGACTCTTCCTTTACTCTTATGCCATTGGTCCCATCATTCTTGTGGGGGAA AAATTCTAGTATTTTGATTATTTGAATCTTACAGCAACAAATAGGAACTCCTGGCCAATGAG AGCTCTTGACCAGTGAATCACCAGCCGATACGAACGTCTTGCCAACAAAATGTGTGGCAAA TAGAAGTATATCAAGCAATAATCTCCCACCCAAGGCTTCTGTAAACTGGGACCAATGATTAC $\tt CTCATAGGGCTGTTGTGAGGATTAGGATGAAATACCTGTGAAAGTGCCTAGGCAGTGCCAGC$ CAAATAGGAGGCATTCAATGAACATTTTTTTGCATATAAACCAAAAAATAACTTGTTATCAAT AAAAACTTGCATCCAACATGAATTTCCAGCCGATGATAATCCAGGCCAAAGGTTTAGTTGTT GTTATTTCCTCTGTATTATTTTCTTCATTACAAAAGAAATGCAAGTTCATTGTAACAATCCA AACAATACCTCACGATATAAAATAAAATGAAAGTATCCTCCTCAAAAA

FIGURE 120

MVAATVAAAWLLLWAAACAQQEQDFYDFKAVNIRGKLVSLEKYRGSVSLVVNVASECGFTDQ HYRALQQLQRDLGPHHFNVLAFPCNQFGQQEPDSNKEIESFARRTYSVSFPMFSKIAVTGTG AHPAFKYLAQTSGKEPTWNFWKYLVAPDGKVVGAWDPTVSVEEVRPQITALVRKLILLKREDL

FIGURE 121

CGGACGCGTGGGCCGGGCCGGGCCAGGCCAAAGCGAGCCATGCCTCTACGTCGGGATGC TGCGCCTGGGGAGCTGTGCGCCGGGAGCTCGGGGGGTGCTGGGGGGCCCGGGCCCTCTCT CGGAGTTGGCAGGAAGCCAGGTTGCAGGGTGTCCGCTTCCTCAGTTCCAGAGAGGTGGATCG CATGGTCTCCACGCCCATCGGAGGCCTCAGCTACGTTCAGGGGTGCACCAAAAAGCATCTTA ACAGCAAGACTGTGGGCCAGTGCCTGGAGACCACAGCACAGAGGGTCCCAGAACGAGAGGCC TTGGTCGTCCTCCATGAAGACGTCAGGTTGACCTTTGCCCAACTCAAGGAGGAGGTGGACAA AGCTGCTTCTGGCCTCCTGAGCATTGGCCTCTGCAAAGGTGACCGGCTGGGCATGTGGGGAC CTAACTCCTATGCATGGGTGCTCATGCAGTTGGCCACCGCCCAGGCGGGCATCATTCTGGTG TCTGTGAACCCAGCCTACCAGGCTATGGAACTGGAGTATGTCCTCAAGAAGGTGGGCTGCAA GGCCCTTGTGTTCCCCAAGCAATTCAAGACCCAGCAATACTACAACGTCCTGAAGCAGATCT GTCCAGAAGTGGAGAATGCCCAGCCAGGGGCCTTGAAGAGTCAGAGGCTCCCAGATCTGACC ACAGTCATCTCGGTGGATGCCCCTTTGCCGGGGACCCTGCTCCTGGATGAAGTGGTGGCGGC TGGCAGCACCAGCATCTGGACCAGCTCCAATACAACCAGCAGTTCCTGTCCTGCCATG ACCCCATCAACATCCAGTTCACCTCGGGGACAACAGGCAGCCCCAAGGGGGCCCACCCTCTCC CACTACAACATTGTCAACAACTCCAACATTTTAGGAGAGCGCCTGAAACTGCATGAGAAGAC ACCAGAGCAGTTGCGGATGATCCTGCCCAACCCCCTGTACCATTGCCTGGGTTCCGTGGCAG GCACAATGATGTCTGATGTACGGTGCCACCCTCATCCTGGCCTCTCCCATCTTCAATGGC AAGAAGGCACTGGAGGCCATCAGCAGAGAGAGAGGCACCTTCCTGTATGGTACCCCCACGAT GTTCGTGGACATTCTGAACCAGCCAGACTTCTCCAGTTATGACATCTCGACCATGTGTGGAG GTGTCATTGCTGGGTCCCCTGCACCTCCAGAGTTGATCCGAGCCATCATCAACAAGATAAAT ATGAAGGACCTGGTGGTTGCTTATGGAACCACAGAGAACAGTCCCGTGACATTCGCGCACTT CCCTGAGGACACTGTGGAGCAGAAGGCAGAAAGCGTGGGCAGAATTATGCCTCACACGGAGG CCCGGATCATGAACATGGAGGCAGGGACGCTGGCAAAGCTGAACACGCCCGGGGAGCTGTGC ATCCGAGGGTACTGCGTCATGCTGGGCTACTGGGGTGAGCCTCAGAAGACAGAGGAAGCAGT AGATCGTGGGCCGCTCTAAGGATATGATCATCCGGGGTGGTGAGAACATCTACCCCGCAGAG CGATCGGATGGGGGAAGAGTTTGTGCCTGCATTCGGCTGAAGGACGGGGAGGAGACCACGG TGGAGGAGATAAAAGCTTTCTGCAAAGGGAAGATCTCTCACTTCAAGATTCCGAAGTACATC GTGTTTGTCACAAACTACCCCCTCACCATTTCAGGAAAGATCCAGAAATTCAAACTTCGAGA $\tt GCAGATGGAACGACATCTAAATCTG{\color{red}{\textbf{TGA}}} ATAAAGCAGCAGGCCTGTCCTGGCCGGTTGGCTT$ GACTCTCTCTGTCAGAATGCAACCTGGCTTTATGCACCTAGATGTCCCCAGCACCCAGTTC TCCATCCCCACATTCCCCTGTCTGTCTTTTTTGGCATAAAGAGCTTCTGTTTTCTTT GAAAAAAAAAAAA

FIGURE 122

MAVYVGMLRLGRLCAGSSGVLGARAALSRSWQEARLQGVRFLSSREVDRMVSTPIGGLSYVQ
GCTKKHLNSKTVGQCLETTAQRVPEREALVVLHEDVRLTFAQLKEEVDKAASGLLSIGLCKG
DRLGMWGPNSYAWVLMQLATAQAGIILVSVNPAYQAMELEYVLKKVGCKALVFPKQFKTQQY
YNVLKQICPEVENAQPGALKSQRLPDLTTVISVDAPLPGTLLLDEVVAAGSTRQHLDQLQYN
QQFLSCHDPINIQFTSGTTGSPKGATLSHYNIVNNSNILGERLKLHEKTPEQLRMILPNPLY
HCLGSVAGTMMCLMYGATLILASPIFNGKKALEAISRERGTFLYGTPTMFVDILNQPDFSSY
DISTMCGGVIAGSPAPPELIRAIINKINMKDLVVAYGTTENSPVTFAHFPEDTVEQKAESVG
RIMPHTEARIMNMEAGTLAKLNTPGELCIRGYCVMLGYWGEPQKTEEAVDQDKWYWTGDVAT
MNEQGFCKIVGRSKDMIIRGGENIYPAELEDFFHTHPKVQEVQVVGVKDDRMGEEICACIRL
KDGEETTVEEIKAFCKGKISHFKIPKYIVFVTNYPLTISGKIQKFKLREQMERHLNL

Signal Peptide:

amino acids 1-22

Transmembrane Domains:

amino acids 140-161, 213-229, 312-334

Putative AMP-binding Domain Signature:

amino acids 260-271

N-myristoylation Sites:

amino acids 19-24, 22-27, 120-125, 203-208, 268-273, 272-277, 314-319, 318-323, 379-384, 380-385, 409-413

N-glycosylation Site:

amino acids 282-285

FIGURE 123

FIGURE 124

 ${\tt GAGCAGGACCC} \underline{\textbf{ATG}} {\tt GACCCCGCCAGGAAAGCAGGTGCCCAGGCCATGATCTGGACTGC}$ AGGCTGGCTGCTGCTGCTTCGCGGAGGGCGCAGGCCCTGGAGTGCTACAGCTGCG TGCAGAAAGCAGATGACGGATGCTCCCCGAACAAGATGAAGACAGTGAAGTGCGCGCCGGGC GTGGACGTCTGCACCGAGGCCGTGGGGGCGGTGGAGACCATCCACGGACAATTCTCGCTGGC AGTGCGGGGTTGCGGTCCCCGGCAAGAATGACCGCGGCCTGGATCTTCACGGGC TTCTGGCGTTCATCCAGCTGCAGCAATGCGCTCAGGATCGCTGCAACGCCAAGCTCAACCTC ACCTCGCGGGCGCTCGACCCGGCAGGTAATGAGAGTGCATACCCGCCCAACGGCGTGGAGTG CTACAGCTGTGTGGGCCTGAGCCGGGAGGCGTGCCAGGGTACATCGCCGCCGGTCGTGAGCT GCTACAACGCCAGCGATCATGTCTACAAGGGCTGCTTCGACGGCAACGTCACCTTGACGGCA GCTAATGTGACTGTCCTTGCCTGTCCGGGGCTGTCCAGGATGAATTCTGCACTCGGGA TGGAGTAACAGGCCCAGGGTTCACGCTCAGTGGCTCCTGTTGCCAGGGGTCCCGCTGTAACT CTGACCTCCGCAACAAGACCTACTTCTCCCCTCGAATCCCACCCCTTGTCCGGCTGCCCCCT CCAGAGCCCACGACTGTGGCCTCAACCACATCTGTCACCACTTCTACCTCGGCCCCAGTGAG ACCCACATCCACCAAACCCATGCCAGCGCCAACCAGTCAGACTCCGAGACAGGGAGTAG AACACGAGGCCTCCCGGGATGAGGAGCCCAGGTTGACTGGAGGCGCCGCTGGCCACCAGGAC CGCAGCAATTCAGGGCAGTATCCTGCAAAAGGGGGGCCCCAGCAGCCCCATAATAAAGGCTG ${\tt TGTGGCTCCCACAGCTGGATTGGCAGCCCTTCTGTTGGCCGTGGCTGCTGGTGTCCTACTG} \underline{{\tt T}}$ **GA**GCTTCTCCACCTGGAAATTTCCCTCTCACCTACTTCTCTGGCCCTGGGTACCCCTCTTCT CATCACTTCCTGTTCCCACCACTGGACTGGGCTGGCCCAGCCCCTGTTTTTCCAACATTCCC CAGTATCCCCAGCTTCTGCTGCGCTGGTTTGCGGCTTTTGGGAAATAAAATACCGTTGTATAT ATTCTGCCAGGGGTGTTCTAGCTTTTTGAGGACAGCTCCTGTATCCTTCTCATCCTTGTCTC TCCGCTTGTCCTCTTGTGATGTTAGGACAGAGTGAGAGAGTCAGCTGTCACGGGGAAGGTG GGTGGGTGGGACAATGGCTCCCCACTCTAAGCACTGCCTCCCCTACTCCCCGCATCTTTGGG GAATCGGTTCCCCATATGTCTTCCTTACTAGACTGTGAGCTCCTCGAGGGGGGGCCCGGTAC CCAATTCGCCCTATAGTGAGTCGTA

FIGURE 125

MDPARKAGAQAMIWTAGWLLLLLRGGAQALECYSCVQKADDGCSPNKMKTVKCAPGVDVCT EAVGAVETIHGQFSLAVRGCGSGLPGKNDRGLDLHGLLAFIQLQQCAQDRCNAKLNLTSRAL DPAGNESAYPPNGVECYSCVGLSREACQGTSPPVVSCYNASDHVYKGCFDGNVTLTAANVTV SLPVRGCVQDEFCTRDGVTGPGFTLSGSCCQGSRCNSDLRNKTYFSPRIPPLVRLPPPEPTT VASTTSVTTSTSAPVRPTSTTKPMPAPTSQTPRQGVEHEASRDEEPRLTGGAAGHQDRSNSG QYPAKGGPQQPHNKGCVAPTAGLAALLLAVAAGVLL

FIGURE 126

 $\tt CGGGACTCGGCGGTCCTCCTGGGAGTCTCGGAGGGGACCGGCTGTGCAGACGCC{\color{red} {\bf ATG}}{\bf GAGT}$ TGGTGCTGGTCTTCCTCTGCAGCCTGCTGGCCCCCATGGTCCTGGCCAGTGCAGCTGAAAAG GAGAAGGAAATGGACCCTTTTCATTATGATTACCAGACCCTGAGGATTGGGGGACTGGTGTT CGCTGTGGTCCTCTTCTCGGTTGGGATCCTCCTTATCCTAAGTCGCAGGTGCAAGTGCAGTT TCAATCAGAAGCCCCGGGCCCCAGGAGATGAGGAAGCCCAGGTGGAGAACCTCATCACCGCC AATGCAACAGAGCCCCAGAAGCAGAGAACTGAAGTGCAGCCATCAGGTGGAAGCCTCTGGAA CCTGAGGCGGCTGCTTGAACCTTTGGATGCAAATGTCGATGCT**TAA**GAAAACCGGCCACTTC AGCAACAGCCCTTTCCCCAGGAGAAGCCAAGAACTTGTGTGTCCCCCACCCTATCCCCTCTA ACACCATTCCTCCACCTGATGATGCAACTAACACTTGCCTCCCCACTGCAGCCTGCGGTCCT CCCAGGCAGGGGCTGAGCCACATGGCCATCTGCTCCTGCCCCGTGGCCCTCCATCAC CTTCTGCTCCTAGGAGGCTGCTTGTTGCCCGAGACCAGCCCCCTCCCCTGATTTAGGGATGC GTAGGGTAAGAGCACGGGCAGTGGTCTTCAGTCGTCTTGGGGACCTGGGAAGGTTTGCAGCAC TTTGTCATCATTCTTCATGGACTCCTTTCACTCCTTTAACAAAACCTTGCTTCCTTATCCC ACCTGATCCCAGTCTGAAGGTCTCTTAGCAACTGGAGATACAAAGCAAGGAGCTGGTGAGCC CAGCGTTGACGTCAGGCAGGCTATGCCCTTCCGTGGTTAATTTCTTCCCAGGGGCTTCCACG AGGAGTCCCCATCTGCCCCGCCCCTTCACAGAGCGCCCGGGGATTCCAGGCCCAGGGCTTCT ACTCTGCCCCTGGGAATGTGTCCCCTGCATATCTTCTCAGCAATAACTCCATGGGCTCTGG GACCCTACCCCTTCCAACCTTCCCTGCTTCTGAGACTTCAATCTACAGCCCAGCTCATCCAG GTTGGGGCCAGCACCCGGGATGGATGGAGGGGAGAGCAGAGGCCTTTGCTTCTCTGCCTACG TCCCCTTAGATGGGCAGCAGGCAACTCCCGCATCCTTTGCTCTGCCTGTCGGTGGTCAGA GCGGTGAGCGAGGTTGGAGACTCAGCAGGCTCCGTGCAGCCCTTGGGAACAGTGAGAG GTTGAAGGTCATAACGAGAGTGGGAACTCAACCCAGATCCCGCCCCTCCTGTCCTGTGTT CCCGCGGAAACCAACCAAACCGTGCGCTGTGACCCATTGCTGTTCTCTGTATCGTGATCTAT CCTCAACAACAACAGAAAAAAGGAATAAAATATCCTTTGTTTCCT

FIGURE 127

 ${\tt MELVLVFLCSLLAPMVLASAAEKEKEMDPFHYDYQTLRIGGLVFAVVLFSVGILLILSRRCK} \\ {\tt CSFNQKPRAPGDEEAQVENLITANATEPQKQRTEVQPSGGSLWNLRRLLEPLDANVDA}$

FIGURE 128

FIGURE 129

 ${\tt MKIPVLPAVVLLSLLVLHSAQGATLGGPEEESTIENYASRPEAFNTPFLNIDKLRSAFKADE}$ ${\tt FLNWHALFESIKRKLPFLNWDAFPKLKGLRSATPDAQ}$

FIGURE 130

FIGURE 131

 ${\tt MGVEIAFASVILTCLSLLAAGVSQVVLLQPVPTQETGPKAMGDLSCGFAGHS}$

FIGURE 132

 ${\tt GGGGAATCTGCAGTAGGTCTGCCGGCG} \underline{\textbf{ATG}} {\tt GAGTGGTGGGCTAGCTCGCCGCTTCGGCTCTG}$ GCTGCTGTTGTTCCTCCTGCCCTCAGCGCAGGGCCCCAGAAGGAGTCAGGTTCAAAATGGA AAGTATTTATTGACCAAATTAACAGGTCTTTGGAGAATTACGAACCATGTTCAAGTCAAAAC TGCAGCTGCTACCATGGTGTCATAGAAGAGGGTCTAACTCCTTTCCGAGGAGGCATCTCCAG GAAGATGATGGCAGAGGTAGTCAGACGGAAGCTAGGGACCCACTATCAGATCACTAAGAACA GACTGTACCGGGAAAATGACTGCATGTTCCCCTCAAGGTGTAGTGGTGTTGAGCACTTTATT TTGGAAGTGATCGGGCGTCTCCCTGACATGGAGATGGTGATCAATGTACGAGATTATCCTCA GGTTCCTAAATGGATGGAGCCTGCCATCCCAGTCTTCTCCTTCAGTAAGACATCAGAGTACC ATGATATCATGTATCCTGCTTGGACATTTTGGGAAGGGGGACCTGCTGTTTGGCCAATTTAT CCTACAGGTCTTGGACGGTGGGACCTCTTCAGAGAAGATCTGGTAAGGTCAGCAGCACAGTG GCCATGGAAAAAGAAAACTCTACAGCATATTTCCGAGGATCAAGGACAAGTCCAGAACGAG ATCCTCTCATTCTTCTGTCTCGGAAAAACCCAAAACTTGTTGATGCAGAATACACCAAAAAC CAGGCCTGGAAATCTATGAAAGATACCTTAGGAAAGCCAGCTGCTAAGGATGTCCATCTTGT GGATCACTGCAAATACAAGTATCTGTTTAATTTTCGAGGCGTAGCTGCAAGTTTCCGGTTTA AACACCTCTTCCTGTGTGGCTCACTTGTTTTCCATGTTGGTGATGAGTGGCTAGAATTCTTC TATCCACAGCTGAAGCCATGGGTTCACTATATCCCAGTCAAAACAGATCTCTCCAATGTCCA AGAGCTGTTACAATTTGTAAAAGCAAATGATGATGTAGCTCAAGAGATTGCTGAAAGGGGAA GCCAGTTTATTAGGAACCATTTGCAGATGGATGACATCACCTGTTACTGGGAGAACCTCTTG AGTGAATACTCTAAATTCCTGTCTTATAATGTAACGAGAAGGAAAGGTTATGATCAAATTAT TCCCAAAATGTTGAAAACTGAACTA**TAG**TAGTCATCATAGGACCATAGTCCTCTTTGTGGCA ACAGATCTCAGATATCCTACGGTGAGAAGCTTACCATAAGCTTGGCTCCTATACCTTGAATA TCTGCTATCAAGCCAAATACCTGGTTTTCCTTATCATGCTGCACCCAGAGCAACTCTTGAGA AAGATTTAAAATGTGTCTAATACACTGATATGAAGCAGTTCAACTTTTTGGATGAATAAGGA CCAGAAATCGTGAGATGTGGATTTTGAACCCAACTCTACCTTTCATTTTCTTAAGACCAATC ACAGCTTGTGCCTCAGATCATCCACCTGTGTGAGTCCATCACTGTGAAATTGACTGTGTCCA TGTGATGATGCCCTTTGTCCCATTATTTGGAGCAGAAAATTCGTCATTTGGAAGTAGTACAA CTCATTGCTGGAATTGTGAAATTATTCAAGGCGTGATCTCTGTCACTTTATTTTAATGTAGG AAACCCTATGGGGTTTATGAAAAATACTTGGGGATCATTCTCTGAATGGTCTAAGGAAGCGG TAGCCATGCCATGCAATGATGTAGGAGTTCTCTTTTTGTAAAACCATAAACTCTGTTACTCAG GAGGTTTCTATAATGCCACATAGAAAGAGGCCAATTGCATGAGTAATTATTGCAATTGGATT TCAGGTTCCCTTTTTGTGCCTTCATGCCCTACTTCTTAATGCCTCTCTAAAGCCAAA

FIGURE 133

MEWWASSPLRLWLLLFLLPSAQGRQKESGSKWKVFIDQINRSLENYEPCSSQNCSCYHGVIE EDLTPFRGGISRKMMAEVVRRKLGTHYQITKNRLYRENDCMFPSRCSGVEHFILEVIGRLPD MEMVINVRDYPQVPKWMEPAIPVFSFSKTSEYHDIMYPAWTFWEGGPAVWPIYPTGLGRWDL FREDLVRSAAQWPWKKKNSTAYFRGSRTSPERDPLILLSRKNPKLVDAEYTKNQAWKSMKDT LGKPAAKDVHLVDHCKYKYLFNFRGVAASFRFKHLFLCGSLVFHVGDEWLEFFYPQLKPWVH YIPVKTDLSNVQELLQFVKANDDVAQEIAERGSQFIRNHLQMDDITCYWENLLSEYSKFLSY NVTRRKGYDQIIPKMLKTEL

FIGURE 134

CACCCTCCATTTCTCGCC ATG CCCTGCACTGCTCCTGATCCCTGCTCTCTTTCATCCTGGCCTTTGGCACCGGAGTGGAGTTCGTGCGCTTTACCTCCCTTCGGCCACTTCTT CCGCAGCATCCTTGCCCCCCTGGCATGGGATCTGGGGCTCCTGCTTCTATTTGTTGGGCAGC ACAGCCTCATGGCAGCTGAAAGAGTGAAGGCATGGACATCCCGGTACTTTGGGGTCCTTCAG AGGTCACTGTATGTGGCCTGCACTGCCCTGGCCTTGCAGCTGGTGATGCGGTACTGGGAGCC CATACCCAAAGGCCCTGTGTTGTGGGAGGCTCGGGCTGAGCCATGGGCCACCTGGGTGCCGC TCCTCTGCTTTGTGCTCCATGTCATCTCCTGGCTCCTCATCTTTAGCATCCTTCTCGTCTTT GACTATGCTGAGCTCATGGGCCTCAAACAGGTATACTACCATGTGCTGGGGCTGGGCCGAGCC TCTGGCCCTGAAGTCTCCCCGGGCTCTCAGACTCTTCTCCCACCTGCGCCACCCAGTGTGTG TGGAGCTGCTGACAGTGCTGTGGGTGGCCTACCCTGGGCACGGACCGTCTCCTTGCT TTCCTCCTTACCTCTGCGCCTGGCTCACGGGCTTGATCAGCAAGACCTCCGCTACCT $\mathsf{CCGGGCCCAGCTACAAAGAAAACTCCACCTGCTCTCTCGGCCCCAGGATGGGGAGGCAGAGT$ **GA**GGAGCTCACTCTGGTTACAAGCCCTGTTCTTCCTCTCCCACTGAATTCTAAATCCTTAAC ATCCAGGCCCTGGCTGCTTCATGCCAGAGGCCCAAATCCATGGACTGAAGGAGATGCCCCTT CTACTACTTGAGACTTTATTCTCTGGGTCCAGCTCCATACCCTAAATTCTGAGTTTCAGCCA CTGAACTCCAAGGTCCACTTCTCACCAGCAAGGAAGAGTGGGGGTATGGAAGTCATCTGTCCC TTCACTGTTTAGAGCATGACACTCTCCCCCTCAACAGCCTCCTGAGAAGGAAAGGATCTGCC CTGACCACTCCCTGGCACTGTTACTTGCCTCTGCGCCTCAGGGGTCCCCTTCTGCACCGCT GGCTTCCACTCCAAGAAGGTGGACCAGGGTCTGCAAGTTCAACGGTCATAGCTGTCCCTCCA GGCCCCAACCTTGCCTCACCACTCCCGGCCCTAGTCTCTGCACCTCCTTAGGCCCTGCCTCT GGGCTCAGACCCCAACCTAGTCAAGGGGATTCTCCTGCTCTTAACTCGATGACTTGGGGGCTC

FIGURE 135

MAPALLLIPAALASFILAFGTGVEFVRFTSLRPLLGGIPESGGPDARQGWLAALQDRSILAP LAWDLGLLLLFVGQHSLMAAERVKAWTSRYFGVLQRSLYVACTALALQLVMRYWEPIPKGPV LWEARAEPWATWVPLLCFVLHVISWLLIFSILLVFDYAELMGLKQVYYHVLGLGEPLALKSP RALRLFSHLRHPVCVELLTVLWVVPTLGTDRLLLAFLLTLYLGLAHGLDQQDLRYLRAQLQR KLHLLSRPQDGEAE

Signal sequence:

amino acids 1-13

Transmembrane domains:

amino acids 58-76, 99-113, 141-159, 203-222

N-myristoylation sites:

amino acids 37-43, 42-48, 229-235

FIGURE 136

CCGAGCACAGGAGTTGCCTGCGTTTAGGAGGTGGCTGCGTTGTGGGAAAAGCTATCAAGGA AGAAATTGCCAAACCATGTCTTTTTTTCTGTTTTCAGAGTAGTTCACAACAGATCTGAGTGT TTTAATTAAGCATGGAATACAGAAAACAACAAAAAACTTAAGCTTTAATTTCATCTGGAATT TCACGTGGTGCTCCCGACTACTCACCCCGAGTGTAAAGAACCTTCGGCTCGCGTGCTTCTG ${\tt AGCTGCTGTGG} \underline{\textbf{ATG}} {\tt GCCTCGGCTCTCTGGACTGTCCTTCCGAGTAGGATGTCACTGAGATCC}$ $\tt CTCAAATGGAGCCTCCTGCTGTCACTCCTGAGTTTCTTTGTGATGTGGTACCTCAGCCT$ ${\tt TCCCCACTACAATGTGATAGAACGCGTGAACTGGATGTACTTCTATGAGTATGAGCCGATTT}$ ACAGACAAGACTTTCACTTCACACTTCGAGAGCATTCAAACTGCTCTCATCAAAATCCATTT $\tt CTGGTCATTCTGGTGACCTCCCACCCTTCAGATGTGAAAGCCAGGCCAGGCCATTAGAGTTAC$ ${\tt TTGGGGTGAAAAAAGTCTTGGTGGGGATATGAGGTTCTTACATTTTTCTTATTAGGCCAAG}$ AGGCTGAAAAGGAAGACAAAATGTTGGCATTGTCCTTAGAGGATGAACACCTTCTTTATGGT GACATAATCCGACAAGATTTTTTAGACACATATAATAACCTGACCTTGAAAACCATTATGGC TTTTCATCAATACTGGCAATTTAGTGAAGTATCTTTTAAACCTAAACCACTCAGAGAAGTTT TTCACAGGTTATCCTCTAATTGATAATTATTCCTATAGAGGATTTTACCAAAAAACCCATAT TTCTTACCAGGAGTATCCTTTCAAGGTGTTCCCTCCATACTGCAGTGGGTTGGGTTATATAA TGTCCAGAGATTTGGTGCCAAGGATCTATGAAATGATGGGTCACGTAAAACCCATCAAGTTT GAAGATGTTTATGTCGGGATCTGTTTGAATTTATTAAAAGTGAACATTCATATTCCAGAAGA CAGCCCATGGCTTTTCTTCCAAGGAGATCATCACTTTTTTGGCAGGTCATGCTAAGGAACACC ACATGCCATTAT<u>TAA</u>CTTCACATTCTACAAAAAGCCTAGAAGGACAGGATACCTTGTGGAAA GTGTTAAATAAAGTAGGTACTGTGGAAAATTCATGGGGAGGTCAGTGTGCTGGCTTACACTG AACTGAAACTCATGAAAAACCCAGACTGGAGACTGGAGGGTTACACTTGTGATTTATTAGTC AGGCCCTTCAAAGATGATATGTGGAGGAATTAAATATAAAGGAATTGGAGGTTTTTGCTAAA GAAATTAATAGGACCAAACAATTTGGACATGTCATTCTGTAGACTAGAATTTCTTAAAAGGG AACAATGTAGTCACTTGAAGGTTTTGTGTATATCTTATGTGGATTACCAATTTAAAAATATA TGTAGTTCTGTGTCAAAAAACTTCTTCACTGAAGTTATACTGAACAAAATTTTACCTGTTTT TGGTCATTTATAAAGTACTTCAAGATGTTGCAGTATTTCACAGTTATTATTATTAAAATTA CTTCAACTTTGTGTTTTTAAATGTTTTGACGATTTCAATACAAGATAAAAAGGATAGTGAAT CATTCTTTACATGCAAACATTTTCCAGTTACTTAACTGATCAGTTTATTATTGATACATCAC TCCATTAATGTAAAGTCATAGGTCATTATTGCATATCAGTAATCTCTTGGACTTTGTTAAAT ATTTTACTGTGGTAATATAGAGAAGAATTAAAGCAAGAAAATCTGAAAA

FIGURE 137

MASALWTVLPSRMSLRSLKWSLLLLSLLSFFVMWYLSLPHYNVIERVNWMYFYEYEPIYRQD FHFTLREHSNCSHQNPFLVILVTSHPSDVKARQAIRVTWGEKKSWWGYEVLTFFLLGQEAEK EDKMLALSLEDEHLLYGDIIRQDFLDTYNNLTLKTIMAFRWVTEFCPNAKYVMKTDTDVFIN TGNLVKYLLNLNHSEKFFTGYPLIDNYSYRGFYQKTHISYQEYPFKVFPPYCSGLGYIMSRD LVPRIYEMMGHVKPIKFEDVYVGICLNLLKVNIHIPEDTNLFFLYRIHLDVCQLRRVIAAHG FSSKEIITFWQVMLRNTTCHY

FIGURE 138

FIGURE 139

MKFTIVFAGLLGVFLAPALANYNINVNDDNNNAGSGQQSVSVNNEHNVANVDNNNGWDSWNS IWDYGNGFAATRLFQKKTCIVHKMNKEVMPSIQSLDALVKEKKLQGKGPGGPPPKGLMYSVN PNKVDDLSKFGKNIANMCRGIPTYMAEEMQEASLFFYSGTCYTTSVLWIVDISFCGDTVEN

Signal Peptide:

amino acids 1-20

N-myristoylation Sites:

amino acids 67-72, 118-123, 163-168

Flavodoxin protein homology:

amino acids 156-174

FIGURE 140

CATTTCTGAAACTAATCGTGTCAGAATTGACTTTGAAAAGCATTGCTTTTTACAGAAGTATA TTAACTTTTTAGGAGTAATTTCTAGTTTGGATTGTAATATGAAATAATTTAAAAGGGCTTCG CTCATATAGGAAAATCGCATATGGTCCTAGTATTAAATTCTTATTGCTTACTGATTTTTT ${\tt CAAGCTTATAGTTGAAATATTTTTCAGGAATTAC} \underline{\textbf{ATG}} \texttt{AATGACAGTCTTCGAACCAATGTGT}$ TTGTTCGATTTCAACCAGAGACTATAGCATGTGCTTGCATCTACCTTGCAGCTAGAGCACTT CCAGGAAATCTGCATAGAAACACTTAGGCTTTATACCAGAAAAAAGCCAAACTATGAATTAC TGGAAAAAGAAGTAGAAAAAAAGAAAGTAGCCTTACAAGAAGCCAAATTAAAAGCAAAGGGA TTGAATCCGGATGGAACTCCAGCCCTTTCAACCCTGGGTGGATTTTCTCCAGCCTCCAAGCC ATCATCACCAAGAGAAGTAAAAGCTGAAGAGAAATCACCAATCTCCATTAATGTGAAGACAG TCAAAAAAGAACCTGAGGATAGACAACAGGCTTCCAAAAGCCCTTACAATGGTGTAAGAAAA TTCTAGATCACATACTCCAAGAAGACACTATAATAATAGGCGGAGTCGATCTGGAACATACA GCTCGAGATCAAGAAGCAGGTCCCGCAGTCACAGTGAAAGCCCTCGAAGACATCATAATCAT GGTTCTCCTCACCTTAAGGCCAAGCATACCAGAGATGATTTAAAAAGTTCAAACAGACATGG TCATAAAAGGAAAAATCTCGTTCTCGATCTCAGAGCAAGTCTCGGGATCACTCAGATGCAG $\verb|CCAAGAAACACAGGCATGAAAGGGGACATCATAGGGACAGGCGTGAACGATCTCGCTCTTT| \\$ CTGA CTTTCTTTCCTTTGGCTTTTGCCTATCTAGTGTGTGTATGGACTCAATCAAAACATTAAACGCAAACTGATTAGGATTTGATTTCTTGAAACCCTCTA GGTCTCTAGAACACTGAGGACAGTTTCTTTTGAAAAGAACTATGTTAATTTTTTTGCACATT AAAATGCCCTAGCAGTATCTAATTAAAAACCATGGTCAGGTTCAATTGTACTTTATTATAGT ${\tt TGTGTATTGTTATTGCTATAAGAACTGGAGCGTGAATTCTGTAAAAATGTATCTTATTTTT}$ ATACAGATAAAATTGCAGACACTGTTCTATTTAAGTGGTTATTTGTTTAAATGATGGTGAAT ACTTTCTTAACACTGGTTTGTCTGCATGTGTAAAGATTTTTACAAGGAAATAAAATACAAAT CTTGTTTTTCTAAAAAAAAAAAAAAAAAAAAGT

FIGURE 141

MNDSLRTNVFVRFQPETIACACIYLAARALQIPLPTRPHWFLLFGTTEEEIQEICIETLRLY
TRKKPNYELLEKEVEKRKVALQEAKLKAKGLNPDGTPALSTLGGFSPASKPSSPREVKAEEK
SPISINVKTVKKEPEDRQQASKSPYNGVRKDSKRSRNSRSASRSRSRTRSRSRSHTPRRHYN
NRRSRSGTYSSRSRSRSRSRSHSESPRRHHNHGSPHLKAKHTRDDLKSSNRHGHKRKKSRSRSQ
SKSRDHSDAAKKHRHERGHHRDRRERSRSFERSHKSKHHGGSRSGHGRHRR

FIGURE 142

FIGURE 143

GGTCCTCGGAGTACAGCTGGTGACCCTGCTCACTGCCACCCTCATGCACAGGCTGGCGC CACACTGCTCCTTCGCGCGCTGGCTGCTCTGTAACGGCAGTTTGTTCCGATACAAGCACCCG TCTGAGGAGGAGCTTCGGGCCCTGGCGGGGAAGCCGAGGCCCAGAGGCAGGAAAGAGCGGTG GGCCAATGGCCTTAGTGAGGAGAAGCCACTGTCTGTGCCCCGAGATGCCCCGTTCCAGCTGG AGACCTGCCCCTCACGACCGTGGATGCCCTGGTCCTGCGCTTCTTCCTGGAGTACCAGTGG TTTGTGGACTTTGCTGTGTACTCGGGCGGCGTGTACCTCTTCACAGAGGCCTACTACAT GCTGGGACCAGCCAAGGAGACTAACATTGCTGTTCTGGTGCCTGCTCACGGTGACCTTCT CCATCAAGATGTTCCTGACAGTGACACGGCTGTACTTCAGCGCCGAGGAGGGGGGGTGAGCGC TCTGTCTGCCTCACCTTTGCCTTCCTCTTCCTGCTGCCATGCTGGTGCAAGTGGTGCG GGAGGAGACCCTCGAGCTGGGCCTGGAGCCTGGTCTGGCCAGCATGACCCAGAACTTAGAGC CACTTCTGAAGAAGCAGGGCTGGGACTGGGCGCTTCCTGTGGCCAAGCTGGCTATCCGCGTG GGACTGGCAGTGGTGGGTGCTGGGTGCCTTCCTCACCTTCCCAGGCCTGCGGCTGGC CCAGACCCACCGGGACGCACTGACCATGTCGGAGGACAGACCCATGCTGCAGTTCCTCCTGC ACACCAGCTTCCTGTCTCCCCTGTTCATCCTGTGGCTCTGGACAAAGCCCATTGCACGGGAC TTCCTGCACCAGCCGCTTTGGGGAGACGCGTTTCTCCCTGCTGTCCGATTCTGCCTTCGA CTCTGGGCGCCTCTGGTTGCTGGTGCTGCTGCTGCGGCTGGCGGTGACCCGGCCCC ACCTGCAGGCCTACCTGTGCCTGGCCAAGGCCCGGGTGGAGCAGCTGCGAAGGGAGGCTGGC CGCATCGAAGCCCGTGAAATCCAGCAGAGGGTGGTCCGAGTCTACTGCTATGTGACCGTGGT GAGCTTGCAGTACCTGACGCCGCTCATCCTCACCCTCAACTGCACACTTCTGCTCAAGACGC TGGGAGGCTATTCCTGGGGCCCAGCTCCTCTACTATCCCCCGACCCATCCTCAGCC AGCGCTGCCCCCATCGGCTCTGGGGAGGACGAAGTCCAGCAGACTGCAGCGCGGATTGCCGG GGCCCTGGGTGGCCTACCTCCCCTCTTCCTCCGTGGCGTCCTGGCCTACCTCATCTGGT GGACGGCTGCCAGCTGCTCGCCAGCCTTTTCGGCCTCTACTTCCACCAGCACTTGGCA GGCTCC $\mathsf{\underline{TAG}}$ $\mathsf{CTGCCTGCAGACCCTCCTGGGGCCCTGAGGTCTGTTCCTGGGGCAGCGGGACA$ CTAGCCTGCCCCTCTGTTTGCGCCCCCGTGTCCCCAGCTGCAAGGTGGGGCCGGACTCCCC GGCGTTCCCTTCACCACAGTGCCTGACCCGCGGCCCCCTTGGACGCCGAGTTTCTGCCTCA GAACTGTCTCCTGGGCCCAGCAGCATGAGGGTCCCGAGGCCATTGTCTCCGAAGCGTATG TGCCAGGTTTGAGTGGCGAGGGTGATGCTGGCTGTTCTGAACAAATAAAGGAGCATGCC GATTTTTAA

FIGURE 144

MAVLGVQLVVTLLTATLMHRLAPHCSFARWLLCNGSLFRYKHPSEEELRALAGKPRPRGRKE RWANGLSEEKPLSVPRDAPFQLETCPLTTVDALVLRFFLEYQWFVDFAVYSGGVYLFTEAYY YMLGPAKETNIAVFWCLLTVTFSIKMFLTVTRLYFSAEEGGERSVCLTFAFLFLLLAMLVQV VREETLELGLEPGLASMTQNLEPLLKKQGWDWALPVAKLAIRVGLAVVGSVLGAFLTFPGLR LAQTHRDALTMSEDRPMLQFLLHTSFLSPLFILWLWTKPIARDFLHQPPFGETRFSLLSDSA FDSGRLWLLVVLCLLRLAVTRPHLQAYLCLAKARVEQLRREAGRIEAREIQQRVVRVYCYVT VVSLQYLTPLILTLNCTLLLKTLGGYSWGLGPAPLLSPDPSSASAAPIGSGEDEVQQTAARI AGALGGLLTPLFLRGVLAYLIWWTAACQLLASLFGLYFHQHLAGS

FIGURE 145

FIGURE 146

GGTTCCTACATCCTCATCTGAGAATCAGAGAGCATAATCTTCTTACGGGCCCGTGATTTATTAACGTGGCTT AATCTGAAGGTTCTCAGTCAAATTCTTTGTGATCTACTGATTGTGGGGGGCATGGCAAGGTTTGCTTAAAGGAGC $\tt TTGGCTGGTTTGGGCCCTTGTAGCTGACAGAAGGTGGCCAGGGAGAATGCAGCACACTGCTCGGAGA{\bf ATG} AAGG$ $\hbox{\tt CGCTTCTGTTGCCTGGCTCAGTCCTGCTAACTACATTGACAATGTGGGCAACCTGCACTTCCTG}$ TATTCAGAACTCTGTAAAGGTGCCTCCCACTACGGCCTGACCAAAGATAGGAAGAGGCGCTCACAAGATGGCTG GTGGACTCTGGCCGGAGCAACCGAACTAGGGCCACGGCCCTTTGAGAGATCCACTATTAGAAGCAGATCATTTAA AAAAATAAATCGAGCTTTGAGTGTTCTTCGAAGGACAAAGAGCGGGAGTGCAGTTGCCAACCATGCCGACCAGG GCAGGGAAAATTCTGAAAACACCACTGCCCCTGAAGTCTTTCCAAGGTTGTACCACCTGATTCCAGATGGTGAA ATTACCAGCATCAAGATCAATCGAGTAGATCCCAGTGAAAGCCTCTCTATTAGGCTGGTGGGAGGTAGCGAAAC GAGACATCATTCTAAAGGTCAACGGGATGGACATCAGCAATGTCCCTCACAACTACGCTGTGCGTCTCCTGCGG ${\tt CAGCCCTGCCAGGTGCTGTGGCTGATGCGTGAACAGAAGTTCCGCAGCAGGAACAATGGACAGGCCCCC}$ GGATGCCTACAGACCCCGAGATGACAGCTTTCATGTGATTCTCAACAAAAGTAGCCCCGAGGAGCAGCTTGGAA ${\tt TAAAACTGGTGCGCAAGGTGGATGAGCCTGGGGTTTTCATCTTCAATGTGCTGGATGGCGGTGTGGCATATCGA}$ CATGGTCAGCTTGAGGAGAATGACCGTGTGTTAGCCATCAATGGACATGATCTTCGATATGGCAGCCCAGAAAG TGCGGCTCATCTGATTCAGGCCAGTGAAAGACGTGTTCACCTCGTCGTGTCCCGCCAGGTTCGGCAGCGGAGCC $\tt CTGACATCTTTCAGGAAGCCGGCTGGAACAGCAATGGCAGCTGGTCCCCAGGGCCAGGGGAGAGGAGCAACACT$ $\tt CCCAAGCCCCTCCATCCTACAATTACTTGTCATGAGAAGGTGGTAAATATCCAAAAAGACCCCGGTGAATCTCT$ ${\tt GAGGAGTCATAAGCAGAGATGGAAGAATAAAAACAGGTGACATTTTGTTGAATGTGGATGGGGTCGAACTGACA}$ GAGGTCAGCCGGAGTGAGCAGTGGCATTATTGAAAAGAACATCATCCTCGATAGTACTCAAAGCTTTGGAAGT GTGACTGGTCCCCATCCTGGGTCATGTGGCTGGAATTACCACGGTGCTTGTATAACTGTAAAGATATTGTATTA TTTCATCAAATCCATTGTTGAAGGAACACCAGCATACAATGATGGAAGAATTAGATGTGGTGATATTCTTCTTG $\mathtt{ATTACTCTAACTATTGTTTCTTGGCCTGGCACTTTTTTA}$ AAAAATGTCAGGAAAAGTATGATCATCTAATGAAAGCCAGTTACACCTCAGAAAATATGATTCCAAAAAAATTA ${\tt AAACTACTAGTTTTTTTCAGTGTGGAGGATTTCTCATTACTCTACAACATTGTTTATATTTTTTTCTATTCAAT}$ AAAAAGCCCTAAAACTAAAATGATTGATTTGTATACCCCACTGAATTCAAGCTGATTTAAATTTAAAATTT ${\tt GGTATATGCTGAAGTCTGCCAAGGGTACATTATGGCCATTTTTAATTTACAGCTAAAATATTTTTTAAAATGCA}$ TTGCTGAGAAACGTTGCTTTCATCAAACAAGAATAAATATTTTTCAGAAGTTAAA

FIGURE 147

MKALLLLVLPWLSPANYIDNVGNLHFLYSELCKGASHYGLTKDRKRRSQDGCPDGCASLTAT
APSPEVSAAATISLMTDEPGLDNPAYVSSAEDGQPAISPVDSGRSNRTRARPFERSTIRSRS
FKKINRALSVLRRTKSGSAVANHADQGRENSENTTAPEVFPRLYHLIPDGEITSIKINRVDP
SESLSIRLVGGSETPLVHIIIQHIYRDGVIARDGRLLPGDIILKVNGMDISNVPHNYAVRLL
RQPCQVLWLTVMREQKFRSRNNGQAPDAYRPRDDSFHVILNKSSPEEQLGIKLVRKVDEPGV
FIFNVLDGGVAYRHGQLEENDRVLAINGHDLRYGSPESAAHLIQASERRVHLVVSRQVRQRS
PDIFQEAGWNSNGSWSPGPGERSNTPKPLHPTITCHEKVVNIQKDPGESLGMTVAGGASHRE
WDLPIYVISVEPGGVISRDGRIKTGDILLNVDGVELTEVSRSEAVALLKRTSSSIVLKALEV
KEYEPQEDCSSPAALDSNHNMAPPSDWSPSWVMWLELPRCLYNCKDIVLRRNTAGSLGFCIV
GGYEEYNGNKPFFIKSIVEGTPAYNDGRIRCGDILLAVNGRSTSGMIHACLARLLKELKGRI
TLTIVSWPGTFL

FIGURE 148

FIGURE 149

MKILVAFLVVLTIFGIQSHGYEVFNIISPSNNGGNVQETVTIDNEKNTAIVNIHAGSCSSTT IFDYKHGYIASRVLSRRACFILKMDHQNIPPLNNLQWYIYEKQALDNMFSNKYTWVKYNPLE SLIKDVDWFLLGSPIEKLCKHIPLYKGEVVENTHNVGAGGCAKAGLLGILGISICADIHV

FIGURE 150

ATGGGGCTCCCTGGGCTGTTCTGCTTGGCCGTGCTGCCAGCAGCTTCTCCAAGGCACG GGAGGAAGAATTACCCCTGTGGTCTCCATTGCCTACAAAGTCCTGGAAGTTTTCCCCAAAG GCCGCTGGGTGCTCATAACCTGCTGTGCACCCCAGCCACCACCGCCCATCACCTATTCCCTC TGTGGAACCAAGAACATCAAGGTGGCCAAGAAGGTGGTGAAGACCCACGAGCCGGCCTCCTT CAACCTCAACGTCACACTCAAGTCCAGTCCAGACCTGCTCACCTACTTCTGCCGGGCGTCCT CCACCTCAGGTGCCCATGTGGACAGTGCCAGGCTACAGATGCACTGGGAGCTGTGGTCCAAG CCAGTGTCTGAGCTGCGGGCCAACTTCACTCTGCAGGACAGAGGGGCAGGCCCCAGGGTGGA GATGATCTGCCAGGCGTCCTCGGGCAGCCCACCTATCACCAACAGCCTGATCGGGAAGGATG AGCCAGACATCGGACTGGTTCTGGTGCCAGGCTGCAAACAACGCCAATGTCCAGCACAGCGC CCTCACAGTGGTGCCCCCAGGTGGTGACCAGAAGATGGAGGACTGGCAGGGTCCCCTGGAGA GCCCCATCCTTGCCGCTCTACAGGAGCACCCGCCGTCTGAGTGAAGAGGAGTTTGGG GGGTTCAGGATAGGGAATGGGGAGGTCAGAGGACGCAAAGCAGCCATG**TAG**AATGAACC GTCCAGAGAGCCAAGCACGGCAGAGGACTGCAGGCCATCAGCGTGCACTGTTCGTATTTGGA

FIGURE 151

MGLPGLFCLAVLAASSFSKAREEEITPVVSIAYKVLEVFPKGRWVLITCCAPQPPPPITYSL CGTKNIKVAKKVVKTHEPASFNLNVTLKSSPDLLTYFCRASSTSGAHVDSARLQMHWELWSK PVSELRANFTLQDRGAGPRVEMICQASSGSPPITNSLIGKDGQVHLQQRPCHRQPANFSFLP SQTSDWFWCQAANNANVQHSALTVVPPGGDQKMEDWQGPLESPILALPLYRSTRRLSEEEFG GFRIGNGEVRGRKAAAM

Signal Peptide:

amino acids 1-18

N-glycosylation Sites:

amino acids 86-89, 132-135, 181-184

FIGURE 152

 ${\tt GGTCCTTA} \underline{\textbf{ATG}} {\tt GCAGCAGCCGCCGCTACCAAGATCCTTCTGTGCCTCCCGCTTCTGCTCCTG}$ CTGTCCGGCTGGTCCCGGGCTGGGCGAGCCGACCCTCACTCTTTTGCTATGACATCACCGT CATCCCTAAGTTCAGACCTGGACCACGGTGGTGTGCGGTTCAAGGCCAGGTGGATGAAAAGA CTTTTCTTCACTATGACTGTGGCAACAAGACAGTCACACCTGTCAGTCCCCTGGGGAAGAAA CTAAATGTCACAACGGCCTGGAAAGCACAGAACCCAGTACTGAGAGAGGTGGTGGACATACT TACAGAGCAACTGCGTGACATTCAGCTGGAGAATTACACACCCAAGGAACCCCTCACCCTGC AGGCAAGGATGTCTTGTGAGCAGAAAGCTGAAGGACACAGCAGTGGATCTTGGCAGTTCAGT TTCGATGGGCAGATCTTCCTCTTTTGACTCAGAGAAGAGAATGTGGACAACGGTTCATCC TGGAGCCAGAAAGATGAAAAGTGGGAGAATGACAAGGTTGTGGCCATGTCCTTCCATT ACTTCTCAATGGGAGACTGTATAGGATGGCTTGAGGACTTCTTGATGGGCATGGACAGCACC CTGGAGCCAAGTGCAGGAGCACCACTCGCCATGTCCTCAGGCACAACCCAACTCAGGGCCAC AGCCACCACCTCATCCTTTGCTGCCTCCTCATCATCCTCCCTGCTTCATCCTCCCTGGCA TCTGAGGAGAGTCCTTTAGAGTGACAGGTTAAAGCTGATACCAAAAGGCTCCTGTGAGCACG GTCTTGATCAAACTCGCCCTTCTGTCTGGCCAGCTGCCCACGACCTACGGTGTATGTCCAGT GGCCTCCAGCAGATCATGACATCATGGACCCAATAGCTCATTCACTGCCTTGATTCCTT TTGCCAACAATTTTACCAGCAGTTATACCTAACATATTATGCAATTTTCTCTTGGTGCTACC GTCAGTAAAATAATCACGTTAGACTTCAGACCTCTGGGGGATTCTTTCCGTGTCCTGAAAGAG AATTTTTAAATTATTTAATAAGAAAAATTTATATTAATGATTGTTTCCTTTAGTAATTTAT

FIGURE 153

MAAAAATKILLCLPLLLLSGWSRAGRADPHSLCYDITVIPKFRPGPRWCAVQGQVDEKTFL HYDCGNKTVTPVSPLGKKLNVTTAWKAQNPVLREVVDILTEQLRDIQLENYTPKEPLTLQAR MSCEQKAEGHSSGSWQFSFDGQIFLLFDSEKRMWTTVHPGARKMKEKWENDKVVAMSFHYFS MGDCIGWLEDFLMGMDSTLEPSAGAPLAMSSGTTQLRATATTLILCCLLIILPCFILPGI

Important features:

Signal peptide:

amino acids 1-25

Transmembrane domain:

amino acids 224-246

N-glycosylation site.

amino acids 68-72, 82-86

N-myristoylation site.

amino acids 200-206, 210-216

Amidation site.

amino acids 77-81

FIGURE 154

FIGURE 155

MELIPTITSWRVLILVVALTQFWCGFLCRGFHLQNHELWLLIKREFGFYSKSQYRTWQKKLA EDSTWPPINRTDYSGDGKNGFYINGGYESHEQIPKRKLKLGGQPTEQHFWARL

FIGURE 156

GTTCTCCTTTCCGAGCCAAAATCCCAGGCGATGGTGAATTATGAACGTGCCACACC**ATG**AAG CTCTTGTGGCAGGTAACTGTGCACCACCACCTGGAATGCCATCCTGCTCCCGTTCGTCTA CCTCACGGCGCAAGTGTGGATTCTGTGTGCAGCCATCGCTGCTGCCGCCTCAGCCGGGCCCC AGAACTGCCCCTCCGTTTGCTCGTGCAGTAACCAGTTCAGCAAGGTGGTGTGCACGCGCGG GGCCTCTCCGAGGTCCCGCAGGGTATTCCCTCGAACACCCGGTACCTCAACCTCATGGAGAA CAACATCCAGATGATCCAGGCCGACACCTTCCGCCACCTCCACCACCTGGAGGTCCTGCAGT ACCCTGGAGCTGTTCGACAACTGGCTGACAGTCATCCCTAGCGGGGCCTTTGAATACCTGTC CAAGCTGCGGAGCTCTGGCTTCGCAACACCCCATCGAAAGCATCCCCTCTTACGCCTTCA ACCGGGTGCCCTCCTCATGCGCCTGGACTTGGGGGGAGCTCAAGAAGCTGGAGTATATCTCT GAGGGAGCTTTTGAGGGGCTGTTCAACCTCAAGTATCTGAACTTGGGCATGTGCAACATTAA AGACATGCCCAATCTCACCCCCTGGTGGGGGCTGGAGGAGCTGGAGATGTCAGGGAACCACT TCCCTGAGATCAGGCCTGGCTCCTTCCATGGCCTGAGCTCCCTCAAGAAGCTCTGGGTCATG CAACTTGGCCCACAATAACCTCTCTTTTTGCCCCCATGACCTCTTTACCCCGCTGAGGTACC TGGTGGAGTTGCATCTACACCACAACCCTTGGAACTGTGATTGTGACATTCTGTGGCTAGCC TGGTGGCTTCGAGAGTATATACCCACCAATTCCACCTGCTGTGGCCGCTGTCATGCTCCCAT GCACATGCGAGGCCGCTACCTCGTGGAGGTGGACCAGGCCTCCTTCCAGTGCTCTGCCCCCT TCATCATGGACGCACCTCGAGACCTCAACATTTCTGAGGGTCGGATGGCAGAACTTAAGTGT CGGACTCCCCTATGTCCTCCGTGAAGTGGTTGCTGCCCAATGGGACAGTGCTCAGCCACGC CTCCCGCCACCAAGGATCTCTGTCCTCAACGACGCACCTTGAACTTTTCCCACGTGCTGC TTTCAGACACTGGGGTGTACACATGCATGGTGACCAATGTTGCAGGCAACTCCAACGCCTCG GCCTACCTCAATGTGAGCACGGCTGAGCTTAACACCTCCAACTACAGCTTCTTCACCACAGT AACAGTGGAGACCACGGAGATCTCGCCTGAGGACACAACGCGAAAGTACAAGCCTGTTCCTA CCACGTCCACTGGTTACCAGCCGGCATATACCACCTCTACCACGGTGCTCATTCAGACTACC CGTGTGCCCAAGCAGGTGGCAGTACCCGCGACAGACACCACTGACAAGATGCAGACCAGCCT GGATGAAGTCATGAAGACCACCAAGATCATCATTGGCTGCTTTGTGGCAGTGACTCTGCTAG CTGCCGCCATGTTGATTGTCTTCTATAAACTTCGTAAGCGGCACCAGCAGCGGAGTACAGTC ACAGCCGCCCGGACTGTTGAGATAATCCAGGTGGACGAAGACATCCCAGCAGCAACATCCGC AGCAGCAACAGCTCCGTCCGGTGTATCAGGTGAGGGGGCAGTAGTGCTGCCCACAATTC ATGACCATATTAACTACAACACCTACAAACCAGCACATGGGGCCCACTGGACAGAAAACAGC CTGGGGAACTCTCTGCACCCACAGTCACCACTATCTCTGAACCTTATATAATTCAGACCCA TACCAAGGACAAGGTACAGGAAACTCAAATA**TGA**CTCCCCTCCCCCAAAAAACTTATAAAAT GCAATAGAATGCACACAAAGACAGCAACTTTTGTACAGAGTGGGGAGAGACTTTTTCTTGTA TATGCTTATATATTAAGTCTATGGGCTGGTTAAAAAAAACAGATTATATTAAAATTTAAAGA CAAAAGTCAAAACA

FIGURE 157

MKLLWQVTVHHHTWNAILLPFVYLTAQVWILCAAIAAAASAGPQNCPSVCSCSNQFSKVVCT
RRGLSEVPQGIPSNTRYLNLMENNIQMIQADTFRHLHHLEVLQLGRNSIRQIEVGAFNGLAS
LNTLELFDNWLTVIPSGAFEYLSKLRELWLRNNPIESIPSYAFNRVPSLMRLDLGELKKLEY
ISEGAFEGLFNLKYLNLGMCNIKDMPNLTPLVGLEELEMSGNHFPEIRPGSFHGLSSLKKLW
VMNSQVSLIERNAFDGLASLVELNLAHNNLSSLPHDLFTPLRYLVELHLHHNPWNCDCDILW
LAWWLREYIPTNSTCCGRCHAPMHMRGRYLVEVDQASFQCSAPFIMDAPRDLNISEGRMAEL
KCRTPPMSSVKWLLPNGTVLSHASRHPRISVLNDGTLNFSHVLLSDTGVYTCMVTNVAGNSN
ASAYLNVSTAELNTSNYSFFTTVTVETTEISPEDTTRKYKPVPTTSTGYQPAYTTSTTVLIQ
TTRVPKQVAVPATDTTDKMQTSLDEVMKTTKIIIGCFVAVTLLAAAMLIVFYKLRKRHQQRS
TVTAARTVEIIQVDEDIPAATSAAATAAPSGVSGEGAVVLPTIHDHINYNTYKPAHGAHWTE
NSLGNSLHPTVTTISEPYIIOTHTKDKVOETOI

FIGURE 158

TTCTCATCTCGTCCTTGCCAAGAGAGTACACAGTCATTAATGAAGCCTGCCCTGGAGCAGAGTGGAATATCATG TGTCGGGAGTGCTGTGAATATGATCAGATTGAGTGCGTCTGCCCCGGAAAGAGGGAAGTCGTGGGTTATACCAT GCAAGAGCTGCCGAAATGGCTCATGGGGGGGTACCTTGGATGACTTCTATGTGAAGGGGTTCTACTGTGCAGAG TGCCGAGCAGGCTGGTACGGAGGAGACTGCATGCGATGTGGCCAGGTTCTGCGAGCCCCAAAGGGTCAGATTTT GTTGGAAAGCTATCCCCTAAATGCTCACTGTGAATGGACCATTCATGCTAAACCTGGGTTTGTCATCCAACTAA GATTTGTCATGTTGAGTCTGGAGTTTGACTACATGTGCCAGTATGACTATGTTGAGGTTCGTGATGGAGACAAC CGCGATGGCCAGATCATCAAGCGTGTCTGTGGCAACGAGCGCCAGCTCCTATCCAGAGCATAGGATCCTCACT CCACGTCCTCTCCACTCCGATGGCTCCAAGAATTTTGACGGTTTCCATGCCATTTATGAGGAGATCACAGCAT GCTCCTCATCCCCTTGTTTCCATGACGGCACGTGCGTCCTTGACAAGGCTGGATCTTACAAGTGTGCCTGCTTG TGGGTACCAGAAAATAACAGGGGGCCCTGGGCTTATCAACGGACGCCATGCTAAAAATTGGCACCGTGGTGTCTT TCTTTTGTAACAACTCCTATGTTCTTAGTGGCAATGAGAAAAGAACTTGCCAGCAGAATGGAGAGTGGTCAGGG $\verb|AAACAGCCCATCTGCATAAAAGCCTGCCGAGAACCAAAGATTTCAGACCTGGTGAGAAGGAGAGTTCTTCCGAT|$ CCCCTACCAAGAAGCCAGCCCTTCCCTTTGGAGATCTGCCCATGGGATACCAACATCTGCATACCCAGCTCCAG TATGAGTGCATCTCACCCTTCTACCGCCGCCTGGGCAGCAGCAGGAGGACATGTCTGAGGACTGGGAAGTGGAG TGGGCGGCACCATCCTGCATCCCTATCTGCGGGAAAATTGAGAACATCACTGCTCCAAAGACCCAAGGGTTGC GCTGGCCGTGGCAGCCATCTACAGGAGGACCAGCGGGGTGCATGACGGCAGCCTACACAAGGGAGCGTGG TTCCTAGTCTGCAGCGGTGCCCTGGTGAATGAGCGCACTGTGGTGGTGGCTGCCCACTGTGTTACTGACCTGGG GAAGGTCACCATGATCAAGACAGCAGACCTGAAAGTTGTTTTGGGGAAATTCTACCGGGATGATGACCGGGATG AGAAGACCATCCAGAGCCTACAGATTTCTGCTATCATTCTGCATCCCAACTATGACCCCATCCTGCTTGATGCT GCCCTGGCTTCAAGAACGACACTGCGCTCTGGGGTGGTCAGTGTGGTGGACTCGCTGCTGTGTGAGGAGCAG CATGAGGACCATGGCATCCCAGTGAGTGTCACTGATAACATGTTCTGTGCCAGCTGGGAACCCACTGCCCCTTC $\tt TGATATCTGCACTGCAGAGACAGGAGGCATCGCGGGTGTGTCCTTCCCGGGACGAGCATCTCCTGAGCCACGCT$ $\tt GTGCTGCCTTTTAAAGACTGGATTGAAAGAAATATGAAA<math>\underline{\textbf{TGA}}$ ACCATGCTCATGCACTCCTTGAGAAGTGTTTC ${\tt TGTATATCCGTCTGTACGTGTCATTGCGTGAAGCAGT}\overline{\tt GTG}\overline{\tt GGCCTGAAGTGTGATTTGGCCTGTGAACTTGG}$ $\tt CTGTGCCAGGGCTTCTGACTTCAGGGACAAAACTCAGTGAAGGGTGAGTAGACCTCCATTGCTGGTAGGCTGAT$ ATATACAAAACCTCTCCACTCACTGACCTGGTGGTCTTCCCCAACTTTCAGTTATACGAATGCCATCAGCTTG ACCAGGGAAGATCTGGGCTTCATGAGGCCCCTTTTGAGGCTCTCAAGTTCTAGAGAGCTGCCTGTGGGACAGCC

FIGURE 159

MELGCWTQLGLTFLQLLLISSLPREYTVINEACPGAEWNIMCRECCEYDQIECVCPGKREVV
GYTIPCCRNEENECDSCLIHPGCTIFENCKSCRNGSWGGTLDDFYVKGFYCAECRAGWYGGD
CMRCGQVLRAPKGQILLESYPLNAHCEWTIHAKPGFVIQLRFVMLSLEFDYMCQYDYVEVRD
GDNRDGQIIKRVCGNERPAPIQSIGSSLHVLFHSDGSKNFDGFHAIYEEITACSSSPCFHDG
TCVLDKAGSYKCACLAGYTGQRCENLLEERNCSDPGGPVNGYQKITGGPGLINGRHAKIGTV
VSFFCNNSYVLSGNEKRTCQQNGEWSGKQPICIKACREPKISDLVRRRVLPMQVQSRETPLH
QLYSAAFSKQKLQSAPTKKPALPFGDLPMGYQHLHTQLQYECISPFYRRLGSSRRTCLRTGK
WSGRAPSCIPICGKIENITAPKTQGLRWPWQAAIYRRTSGVHDGSLHKGAWFLVCSGALVNE
RTVVVAAHCVTDLGKVTMIKTADLKVVLGKFYRDDDRDEKTIQSLQISAIILHPNYDPILLD
ADIAILKLLDKARISTRVQPICLAASRDLSTSFQESHITVAGWNVLADVRSPGFKNDTLRSG
VVSVVDSLLCEEQHEDHGIPVSVTDNMFCASWEPTAPSDICTAETGGIAAVSFPGRASPEPR
WHLMGLVSWSYDKTCSHRLSTAFTKVLPFKDWIERNMK

FIGURE 160

ACCAGGCATTGTATCTTCAGTTGTCATCAAGTTCGCAATCAGATTGGAAAAGCTCAACTTGA AGCTTTCTTGCCTGCAGTGAAGCAGAGAGATAGATATTATTCACGTAATAAAAAAC**ATG**GGC TTCAACCTGACTTTCCACCTTTCCTACAAATTCCGATTACTGTTGCTGTTGACTTTGTGCCT GACAGTGGTTGGGTGGCCACCAGTAACTACTTCGTGGGTGCCATTCAAGAGATTCCTAAAG CAAAGGAGTTCATGGCTAATTTCCATAAGACCCTCATTTTGGGGAAGGGAAAAACTCTGACT AATGAAGCATCCACGAAGAAGGTAGAACTTGACAACTGTCCTTCTGTGTCTCCTTACCTCAG AGGCCAGAGCAAGCTCATTTTCAAACCAGATCTCACTTTGGAAGAGGTACAGGCAGAAAATC CCAAAGTGTCCAGAGGCCGGTATCGCCCTCAGGAATGTAAAGCTTTACAGAGGGTCGCCATC CTCGTTCCCCACCGGAACAGAGAAACACCTGATGTACCTGCTGGAACATCTGCATCCCTT CCTGCAGAGGCAGCTGGATTATGGCATCTACGTCATCCACCAGGCTGAAGGTAAAAAGT TTAATCGAGCCAAACTCTTGAATGTGGGCTATCTAGAAGCCCTCAAGGAAGAAATTGGGAC TGCTTTATATTCCACGATGTGGACCTGGTACCCGAGAATGACTTTAACCTTTACAAGTGTGA GGAGCATCCCAAGCATCTGGTGGTTGGCAGGAACAGCACTGGGTACAGGTTACGTTACAGTG GATATTTTGGGGGTGTTACTGCCCTAAGCAGAGAGCAGTTTTTCAAGGTGAATGGATTCTCT AACAACTACTGGGGATGGGGAGGCGAAGACGATGACCTCAGACTCAGGGTTGAGCTCCAAAG AATGAAAATTTCCCGGCCCCTGCCTGAAGTGGGTAAATATACAATGGTCTTCCACACTAGAG ACAAAGGCAATGAGGTGAACGCAGAACGGATGAAGCTCTTACACCAAGTGTCACGAGTCTGG AGAACAGATGGGTTGAGTAGTTCTTATAAATTAGTATCTGTGGAACACAATCCTTTATA ${\tt TATCAACATCACAGTGGATTTCTGGTTTGGTGCA} \underline{{\tt TGA}} {\tt CCCTGGATCTTTTGGTGATGTTTGG}$ AAGAACTGATTCTTTGTTTGCAATAATTTTGGCCTAGAGACTTCAAATAGTAGCACACATTA AGAACCTGTTACAGCTCATTGTTGAGCTGAATTTTTCCTTTTTGTATTTTCTTAGCAGAGCT CCTGGTGATGTAGAGTATAAAACAGTTGTAACAAGACAGCTTTCTTAGTCATTTTGATCATG AGGGTTAAATATTGTAATATGGATACTTGAAGGACTTTATATAAAAGGATGACTCAAAGGAT TATGGGATAAAAGGCCACAGGAAATAAGACTGCTGAATGTCTGAGAGAACCAGAGTTGTTCT CGTCCAAGGTAGAAAGGTACGAAGATACAATACTGTTATTCATTTATCCTGTACAATCATCT GTGAAGTGGTGTCAGGTGAGAAGGCGTCCACAAAAGAGGGGAGAAAAGGCGACGAATCA GGACACAGTGAACTTGGGAATGAAGAGGTAGCAGGAGGGTGGAGTGTCGGCTGCAAAGGCAG CAGTAGCTGAGCTGCTGCAGGTGCTGATAGCCTTCAGGGGAGGACCTGCCCAGGTATGCCT TCCAGTGATGCCCACCAGAGAATACATTCTCTATTAGTTTTTAAAGAGTTTTTTGTAAAATGA TGTCTATCAAATACCTCTGTAGTAAAATGTGAAAAAAGCAAAA

FIGURE 161

MGFNLTFHLSYKFRLLLLTLCLTVVGWATSNYFVGAIQEIPKAKEFMANFHKTLILGKGKT
LTNEASTKKVELDNCPSVSPYLRGQSKLIFKPDLTLEEVQAENPKVSRGRYRPQECKALQRV
AILVPHRNREKHLMYLLEHLHPFLQRQQLDYGIYVIHQAEGKKFNRAKLLNVGYLEALKEEN
WDCFIFHDVDLVPENDFNLYKÇEEHPKHLVVGRNSTGYRLRYSGYFGGVTALSREQFFKVNG
FSNNYWGWGGEDDDLRLRVELQRMKISRPLPEVGKYTMVFHTRDKGNEVNAERMKLLHQVSR
VWRTDGLSSCSYKLVSVEHNPLYINITVDFWFGA

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites:

amino acids 4-7, 220-223 and 335-338

Xylose isomerase proteins:

amino acids 191-201

FIGURE 162

CGTGGGCCGGGGTCGCGCGGGGCTGTGGGCGCCCGGAGGAGCGACCGCCGCAGTTCTC GAGCTCCAGCTGCATTCCCTCCGCGTCCGCCCCACGCTTCTCCCGGCTCCGGGCCCCGCAATG CCCGGCAGGGTGGCCGCAGGCCTGTATGAACTCAATCTCACCACCGATAGCCCTGCCACCA CGGGAGCGGTGGTGACCATCTCGGCCAGCCTGGTGGCCAAGGACAACGGCAGCCTGGCCCTG CCCGCTGACGCCCACCTCTACCGCTTCCACTGGATCCACACCCCGCTGGTGCTTACTGGCAA GATGGAGAAGGGTCTCAGCTCCACCATCCGTGTGGTCGGCCACGTGCCCGGGGAATTCCCGG TCTCTGTCTGGGTCACTGCCGCTGACTGCTGGATGTGCCAGCCTGTGGCCAGGGGCTTTGTG GTCCTCCCCATCACAGAGTTCCTCGTGGGGGGACCTTGTTGTCACCCAGAACACTTCCCTACC GCAACTTCCTCAAGACCGCCTTGTTTCTCTACAGCTGGGACTTCGGGGACCGGGACCCAGATG GTGACTGAAGACTCCGTGGTCTATTATAACTATTCCATCATCGGGACCTTCACCGTGAAGCT CAAAGTGGTGGCGGAGTGGGAAGAGGTGGAGCCGGATGCCACGAGGGCTGTGAAGCAGAAGA CCGGGGACTTCTCCGCCTCGCTGAAGCTGCAGGAAACCCTTCGAGGCATCCAAGTGTTGGGG CCCACCCTAATTCAGACCTTCCAAAAGATGACCGTGACCTTGAACTTCCTGGGGAGCCCTCC TCTGACTGTGTGCTGGCGTCTCAAGCCTGAGTGCCTCCCGCTGGAGGAAGGGGAGTGCCACC CTGTGTCCGTGGCCAGCACAGCGTACAACCTGACCCACACCTTCAGGGACCCTGGGGACTAC TGCTTCAGCATCCGGGCCGAGAATATCATCAGCAAGACACATCAGTACCACAAGATCCAGGT GTGGCCCTCCAGAATCCAGCCGGCTGTCTTTGCTTTCCCATGTGCTACACTTATCACTGTGA TGTTGGCCTTCATCATGTACATGACCCTGCGGAATGCCACTCAGCAAAAGGACATGGTGGAG AACCCGGAGCCACCCTCTGGGGTCAGGTGCTGCCAGATGTGCTGTGGGCCTTTCTTGCT GGAGACTCCATCTGAGTACCTGGAAATTGTTCGTGAGAACCACGGGCTGCTCCCGCCCCTCT ${\tt ATAAGTCTGTCAAAACTTACACCGTG} {\tt TGA} {\tt GCACTCCCCCTCCCCACCCCATCTCAGTGTTAA}$ CTGACTGCTGACTTGGAGTTTCCAGCAGGGTGGTGTGCACCACTGACCAGGAGGGGTTCATT TGCGTGGGGCTGTTGGCCTGGATCATCCATCTGTACAGTTCAGCCACTGCCACAAGCC CCTCCCTCTCTGTCACCCCTGACCCCAGCCATTCACCCATCTGTACAGTCCAGCCACTGACA TAAGCCCCACTCGGTTACCACCCCCTTGACCCCTACCTTTGAAGAGGCCTTCGTGCAGGACT TTGATGCTTGGGGTGTTCCGTGTTGACTCCTAGGTGGGCCTGCCCACTGCCCATTCCT CTCATATTGGCACATCTGCTGTCCATTGGGGGTTCTCAGTTTCCTCCCCCAGACAGCCCTAC CTGTGCCAGAGAGCTAGAAAGGTCATAAAGGGTTAAAAATCCATAACTAAAGGTTGTAC CACACACACAGAAATATAAACACATGCGTCACATGGGCATTTCAGATGATCAGCTCTGTA TCTGGTTAAGTCGGTTGCTGGGATGCACCCTGCACTAGAGCTGAAAGGAAATTTGACCTCCA AGCAGCCCTGACAGGTTCTGGGCCCGGGCCCTCCCTTTGTGCTTTGTCTCTGCAGTTCTTGC GCCCTTTATAAGGCCATCCTAGTCCCTGCTGGCTGGCAGGGGCCTGGATGGGGGGCAGGACT AATACTGAGTGATTGCAGAGTGCTTTATAAATATCACCTTATTTTATCGAAACCCATCTGTG AAACTTTCACTGAGGAAAAGGCCTTGCAGCGGTAGAAGAGGTTGAGTCAAGGCCGGGCGCGG TGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGTGGATCACGAGATCAGGA GATCGAGACCACCCTGGCTAACACGGTGAAACCCCCGTCTCTACTAAAAAAATACAAAAAGTT AGCCGGGCGTGGTGGGTGCCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATG GTGCGAACCCGGGAGCCTGCAGTGAGCCCAGATGGCGCCACTGCACTCCAGCCTGA GTGACAGAGCGAGACTCTGTCTCCA

FIGURE 163

MAQAVWSRLGRILWLACLLPWAPAGVAAGLYELNLTTDSPATTGAVVTISASLVAKDNGSLA
LPADAHLYRFHWIHTPLVLTGKMEKGLSSTIRVVGHVPGEFPVSVWVTAADCWMCQPVARGF
VVLPITEFLVGDLVVTQNTSLPWPSSYLTKTVLKVSFLLHDPSNFLKTALFLYSWDFGDGTQ
MVTEDSVVYYNYSIIGTFTVKLKVVAEWEEVEPDATRAVKQKTGDFSASLKLQETLRGIQVL
GPTLIQTFQKMTVTLNFLGSPPLTVCWRLKPECLPLEEGECHPVSVASTAYNLTHTFRDPGD
YCFSIRAENIISKTHQYHKIQVWPSRIQPAVFAFPCATLITVMLAFIMYMTLRNATQQKDMV
ENPEPPSGVRCCCQMCCGPFLLETPSEYLEIVRENHGLLPPLYKSVKTYTV

Important features of the protein:

Signal peptide:

amino acids 1-24

Transmembrane domain:

amino acids 339-362

N-glycosylation sites.

amino acids 34-37, 58-61, 142-145, 197-200, 300-303 and 364-367

FIGURE 164

FIGURE 165

 ${\tt MALSSQIWAACLLLLLLASLTSGSVFPQQTGQLAELQPQDRAGARASWMPMFQRRRRRDTH} \\ {\tt FPICIFCCGCCHRSKCGMCCKT}$

FIGURE 166

CTGTCAGGAAGGACCATCTGAAGGCTGCAATTTGTTCTTAGGGAGGCAGGTGCTGGCCTGGC $\tt CTGGATCTTCCACC \underline{\textbf{ATG}} \texttt{TTCCTGTTGCTGCCTTTTGATAGCCTGATTGTCAACCTTCTGGGC$ AGTCTCCTTTGGTATCCGCAAACTCTACATGAAAAGTCTGTTAAAAATCTTTGCGTGGGCTA CCTTGAGAATGGAGCGAGGAGCCAAGGAGAAGAACCACCAGCTTTACAAGCCCTACACCAAC GGAATCATTGCAAAGGATCCCACTTCACTAGAAGAAGAAGATCAAAGAGATTCGTCGAAGTGG TAGTAGTAAGGCTCTGGACAACACTCCAGAGTTCGAGCTCTCTGACATTTTCTACTTTTGCC GGAAAGGAATGGAGACCATTATGGATGATGAGGTGACAAAGAGATTCTCAGCAGAAGAACTG GAGTCCTGGAACCTGCTGAGCAGAACCAATTATAACTTCCAGTACATCAGCCTTCGGCTCAC GGTCCTGTGGGGGTTAGGAGTGCTGATTCGGTACTGCTTCTGCTGCCGCTCAGGATAGCAC TGGCTTTCACAGGGATTAGCCTTCTGGTGGTGGGCACAACTGTGGTGGGATACTTGCCAAAT AGCGCTGACAGCCATCACCATCACCATGACAGGGAAAACAGACCAAGAAATGGTGGCATCT GTGTGGCCAATCATACCTCACCGATCGATGTGATCATCTTGGCCAGCGATGGCTATTATGCC ATGGTGGGTCAAGTGCACGGGGGACTCATGGGTGTGATTCAGAGAGCCATGGTGAAGGCCTG CCCACACGTCTGGTTTGAGCGCTCGGAAGTGAAGGATCGCCACCTGGTGGCTAAGAGACTGA CTGAACATGTGCAAGATAAAAGCAAGCTGCCTATCCTCATCTTCCCAGAAGGAACCTGCATC AATAATACATCGGTGATGATGTTCAAAAAGGGAAGTTTTGAAATTGGAGCCACAGTTTACCC TGTTGCTATCAAGTATGACCCTCAATTTGGCGATGCCTTCTGGAACAGCAGCAAATACGGGA TGGTGACGTACCTGCGAATGATGACCAGCTGGGCCATTGTCTGCAGCGTGTGGTACCTG CCTCCCATGACTAGAGAGGCAGATGAAGATGCTGTCCAGTTTGCGAATAGGGTGAAATCTGC CATTGCCAGGCAGGAGGACTTGTGGACCTGCTGTGGGATGGGGGCCTGAAGAGGGAGAAGG TGAAGGACACGTTCAAGGAGGAGCAGCAGAAGCTGTACAGCAAGATGATCGTGGGGAACCAC ${ t AAGGACAGGAGCCGCTCC} { t TGA} { t GCCTCCAGCTGGCTGGGGCCACCGTGCGGGGTGCCAA}$ CGGGCTCAGAGCTGGAGTTGCCGCCGCCCCCCCCCCTGTGTCCTTTCCAGACTCCAGGG CTCCCCGGGCTGCTCTGGATCCCAGGACTCCGGCTTTCGCCGAGCCGCAGCGGGATCCCTGT GCACCCGGCGCAGCCTACCCTTGGTGGTCTAAACGGATGCTGCTGGGTGTTGCGACCCAGGA CGAGATGCCTTGTTTCTTTTACAATAAGTCGTTGGAGGAATGCCATTAAAGTGAACTCCCCA CCTTTGCACGCTGTGCGGGCTGAGTGGTTGGGGGAGATGTGGCCATGGTCTTGTGCTAGAGAT GGCGGTACAAGAGTCTGTTATGCAAGCCCGTGTGCCAGGGATGTGCTGGGGGGCGGCCACCCG CTCTCCAGGAAAGGCACAGCTGAGGCACTGTGGCTGGCTTCGGCCTCAACATCGCCCCCAGC CTTGGAGCTCTGCAGACATGATAGGAAGGAAACTGTCATCTGCAGGGGCTTTCAGCAAAATG GGCCGCTGACTGGGGCCATGGGGAGAACGTGTTTCGTACTCCAGGCTAACCCTGAACTCCCC ATGTGATGĆGCGCTTTGTTGAATGTGTGTCTCGGTTTCCCCATCTGTAATATGAGTCGGGGG AGGACACATCACGTTCAGTGTTTCAAGTACAGGCCCACAAAACGGGGCACGGCAGGCCTGAG TGA

FIGURE 167

MFLLLPFDSLIVNLLGISLTVLFTLLLVFIIVPAIFGVSFGIRKLYMKSLLKIFAWATLRME RGAKEKNHQLYKPYTNGIIAKDPTSLEEEIKEIRRSGSSKALDNTPEFELSDIFYFCRKGME TIMDDEVTKRFSAEELESWNLLSRTNYNFQYISLRLTVLWGLGVLIRYCFLLPLRIALAFTG ISLLVVGTTVVGYLPNGRFKEFMSKHVHLMCYRICVRALTAIITYHDRENRPRNGGICVANH TSPIDVIILASDGYYAMVGQVHGGLMGVIQRAMVKACPHVWFERSEVKDRHLVAKRLTEHVQ DKSKLPILIFPEGTCINNTSVMMFKKGSFEIGATVYPVAIKYDPQFGDAFWNSSKYGMVTYL LRMMTSWAIVCSVWYLPPMTREADEDAVQFANRVKSAIARQGGLVDLLWDGGLKREKVKDTF KEEQQKLYSKMIVGNHKDRSRS

FIGURE 168

GCCCCTCGAAACCAGGACTCCAGCACCTCTGGTCCCGCCCTCACCCGGACCCCTGGCCCTCA ACCTGGCAGGCCCAGGCTGTTCCCACCATCCTGCCCCTGGGCCTGGCTCCAGACACCTTTGA CGATACCTATGTGGGTTGTGCAGAGGAGATGGAGGAAGGCAGCCCCCCTGCTAAAGGAGG AAATGGCCCACCATGCCTGCTGCGGGAATCCTGGGAGGCCAGGAGACCTGGGAGGAC AAGCGTCGAGGGCTTACCTTGCCCCCTGGCTTCAAAGCCCAGAATGGAATAGCCATTATGGT CTACACCAACTCATCGAACACCTTGTACTGGGAGTTGAATCAGGCCGTGCGGACGGGCGGAG GCTCCCGGGAGCTCTACATGAGGCACTTTCCCTTCAAGGCCCTGCATTTCTACCTGATCCGG GCCCTGCAGCTGCGAGGCAGTGGGGGGCTGCAGCAGGGGACCTGGGGAGGTGGTTCCG AGGTGTGGGCAGCCTTCGCTTTGAACCCAAGAGGCTGGGGGACTCTGTCCGCTTGGGCCAGT TTGCCTCCAGCTCCCTGGATAAGGCAGTGGCCCACAGATTTGGGGAGAAGAGGCGGGGCTGT GTGTCTGCGCCAGGGGTGCAGCTAGGGTCACAATCTGAGGGGGCCTCCTCTCTGCCCCCCTG ${\tt GAAGACTCTGCTCTTGGCCCCTGGAGAGTTCCAGCTCTCAGGGGGTTGGGCCC}{{\tt TGA}}{\tt AAGTCCA}$ ACATCTGCCACTTAGGAGCCCTGGGAACGGGTGACCTTCATATGACGAAGAGGCACCTCCAG CAGCCTTGAGAAGCAAGAACATGGTTCCGGACCCAGCCCTAGCAGCCTTCTCCCCAACCAGG ATGTTGGCCTGGGGAGGCCACAGCAGGGCTGAGGGAACTCTGCTATGTGATGGGGACTTCCT TGGAGTTTTATTGAGGTAGCTACGTGATTAAATGGTATTGCAGTGTGGA

FIGURE 169

MALAALMIALGSLGLHTWQAQAVPTILPLGLAPDTFDDTYVGCAEEMEEKAAPLLKEEMAHH ALLRESWEAAQETWEDKRRGLTLPPGFKAQNGIAIMVYTNSSNTLYWELNQAVRTGGGSREL YMRHFPFKALHFYLIRALQLLRGSGGCSRGPGEVVFRGVGSLRFEPKRLGDSVRLGQFASSS LDKAVAHRFGEKRRGCVSAPGVQLGSQSEGASSLPPWKTLLLAPGEFQLSGVGP

FIGURE 170

GTGGCTTCATTTCAGTGGCTGACTTCCAGAGAGCAAT**ATG**GCTGGTTCCCCAACATGCCTCA CCCTCATCTATATCCTTTGGCAGCTCACAGGGTCAGCAGCCTCTGGACCCGTGAAAGAGCTG GTCGGTTCCGTTGGTGGGCCGTGACTTTCCCCCTGAAGTCCAAAGTAAAGCAAGTTGACTC TATTGTCTGGACCTTCAACACACCCTCTTGTCACCATACAGCCAGAAGGGGGCACTATCA TAGTGACCCAAAATCGTAATAGGGAGAGAGTAGACTTCCCAGATGGAGGCTACTCCCTGAAG CTCAGCAAACTGAAGAAGAATGACTCAGGGATCTACTATGTGGGGGATATACAGCTCATCACT CCAGCAGCCCTCCACCCAGGAGTACGTGCTGCATGTCTACGAGCACCTGTCAAAGCCTAAAG TCACCATGGGTCTGCAGAGCAATAAGAATGGCACCTGTGTGACCAATCTGACATGCTGCATG GAACATGGGGAAGAGGTGTGATTTATACCTGGAAGGCCCTGGGGCAAGCAGCCAATGAGTC CCATAATGGGTCCATCCCCCATCTCCTGGAGATGGGGAGAAAGTGATATGACCTTCATCT GCGTTGCCAGGAACCTGTCAGCAGAAACTTCTCAAGCCCCATCCTTGCCAGGAAGCTCTGT GAAGGTGCTGATGACCCAGATTCCTCCATGGTCCTCTGTGTCTCCTGTTGGTGCCCCT AGTACATTGAAGAGAGAGAGAGTGGACATTTGTCGGGAAACTCCTAACATATGCCCCCAT TCTGGAGAGAACACAGAGTACGACACAATCCCTCACACTAATAGAACAATCCTAAAGGAAGA TCCAGCAAATACGGTTTACTCCACTGTGGAAATACCGAAAAAGATGGAAAATCCCCACTCAC ${\tt TGCTCACGATGCCAGACACCAAGGCTATTTGCCTATGAGAATGTTATC} \underline{{\tt TAG}}{\tt ACAGCAGTG}$ CACTCCCCTAAGTCTCTGCTCA

FIGURE 171

MAGSPTCLTLIYILWQLTGSAASGPVKELVGSVGGAVTFPLKSKVKQVDSIVWTFNTTPLVT
IQPEGGTIIVTQNRNRERVDFPDGGYSLKLSKLKKNDSGIYYVGIYSSSLQQPSTQEYVLHV
YEHLSKPKVTMGLQSNKNGTCVTNLTCCMEHGEEDVIYTWKALGQAANESHNGSILPISWRW
GESDMTFICVARNPVSRNFSSPILARKLCEGAADDPDSSMVLLCLLLVPLLLSLFVLGLFLW
FLKRERQEEYIEEKKRVDICRETPNICPHSGENTEYDTIPHTNRTILKEDPANTVYSTVEIP
KKMENPHSLLTMPDTPRLFAYENVI

FIGURE 172

TCTGGACCCGTGAAAGAGCTGGTCGGTTCCGTTGGTGGGGCCCGTGACTTTCCCCCTGAAGTC CAAAGTAAAGCAAGTTGACTCTATTGTCTGGACCTTCAACACAACCCCTCTTGTCACCATAC AGCCAGAAGGGGGCACTATCATAGTGACCCAAAATCGTAATAGGGAGAGAGTAGACTTCCCA GATGGAGGCTACTCCCTGAAGCTCAGCAAACTGAAGAAGAATGACTCAGGGATCTACTATGT GGGGATATACAGCTCATCACTCCAGCAGCCCTCCACCCAGGAGTACGTGCTGCATGTCTACG AGCACCTGTCAAAGCCTAAAGTCACCATGGGTCTGCAGAGCAATAAGAATGGCACCTGTGTG ACCAATCTGACATGCTGCATGGAACATGGGGAAGAGGATGTGATTTATACCTGGAAGGCCCT GGGGCAAGCAATGAGTCCCATAATGGGTCCATCCTCCCCATCTCCTGGAGATGGGGAG AAAGTGATATGACCTTCATCTGCGTTGCCAGGAACCCTGTCAGCAGAAACTTCTCAAGCCCC ATCCTTGCCAGGAAGCTCTGTGAAGGTGCTGCTGATGACCCAGATTCCTCCATGGTCCTCCT GTGTCTCCTGTTGGTGCCCCTCCTGCTCAGTCTCTTTGTACTGGGGCTATTTCTTTGGTTTC TGAAGAGAGAGAGAAGAAGAGTACATTGAAGAGAAGAAGAGTGGACATTTGTCGGGAA ACTCCTAACATATGCCCCCATTCTGGAGAGAACACAGAGTACGACAATCCCTCACACTAA TAGAACAATCCTAAAGGAAGATCCAGCAAATACGGTTTACTCCACTGTGGAAATACCGAAAA AGATGGAAAATCCCCACTCACTGCTCACGATGCCAGACACCAAGGCTATTTGCCTATGAG

FIGURE 173

GAAAGACGTGGTCCTGACAGACAGACAATCCTATTCCCTACCAAAATGAAGATGCTGCTGCT
GCTGTGTTTGGGACTGACCCTAGTCTGTGTCCATGCAGAAGAAGCTAGTTCTACGGGAAGGA
ACTTTAATGTAGAAAAGATTAATGGGGAATGGCATACTATTATCCTGGCCTCTGACAAAAGA
GAAAAGATAGAAGAACATGGCAACTTTAGACTTTTTCTGGAGCAAATCCATGTCTTGGAGAA
TTCCTTAGTTCTTAAAGTCCATACTGTAAGAGATGAAGAGTGCTCCGAATTATCTATGGTTG
CTGACAAAACAGAAAAGGCTGGTGAATATTCTGTGACGTATGATGGATTCAATACATTTACT
ATACCTAAGACAGACTATGATAACTTTCTTATGGCTCACCTCATTAACGAAAAGGATGGGGA
AACCTTCCAGCTGATGGGGCTCTATGGCCGAGAACCAGATTTGAGTTCAGACATCAAGGAAA
GGTTTGCACAACTATGTGAGGAGCATGGAATCCTTAGAGAAAATATCATTGACCTATCCAAT
GCCAATCGCTGCCTCCAGGCCCGAGAATGAAGAATGGCCTGAGCCTCCAGTGTTGAGTGGAC
ACTTCTCACCAGGACTCCACCATCATCCCTTCCTATCCATACAGCATCCCAGTATAAAATTC
TGTGATCTGCATTCCATCCTGTCTCACTGAGAAGTCCAATTCCAGTCTTATCAACATGTTACC
TAGGATACCTCATCAAGAATCAAAGACTTCTTTAAATTTCTCTTTGATACACCCTTGACAAT
TTTTCATGAAATTATTCCTCTTCCTGTTCAATAAATGATTACCCTTTGCACTTAA

FIGURE 174

MKMLLLLCLGLTLVCVHAEEASSTGRNFNVEKINGEWHTIILASDKREKIEEHGNFRLFLEQ IHVLENSLVLKVHTVRDEECSELSMVADKTEKAGEYSVTYDGFNTFTIPKTDYDNFLMAHLI NEKDGETFQLMGLYGREPDLSSDIKERFAQLCEEHGILRENIIDLSNANRCLQARE

FIGURE 175

FIGURE 176

MTCCEGWTSCNGFSLLVLLLLGVVLNAIPLIVSLVEEDQFSQNPISCFEWWFPGIIGAGLMA IPATTMSLTARKRACCNNRTGMFLSSFFSVITVIGALYCMLISIQALLKGPLMCNSPSNSNA NCEFSLKNISDIHPESFNLQWFFNDSCAPPTGFNKPTSNDTMASGWRASSFHFDSEENKHRL IHFSVFLGLLLVGILEVLFGLSQIVIGFLGCLCGVSKRRSQIV

FIGURE 177

GTCGAATCCAAATCACTCATTGTGAAAGCTGAGCTCACAGCCGAATAAGCCACCATGCTCGCTGCTCTGTCT
GTCAGTGTGTCTCCTGATGGTCTCGCTGGCCCTTTGCTGCTACCAGGCCCATGCTCTTGTCT
GCCCAGCTGTTGCTTCTGAGATCACAGTCTTCTTATTCTTAAGTGACGCTGCGGTAAACCTC
CAAGTTGCCAAACTTAATCCACCTCCAGAAGCTCTTGCAGCCAAGTTGGAAGTGAAGCACTG
CACCGATCAGATATCTTTTAAGAAACGACTCTCATTGAAAAAGTCCTGGTGGAAATAGTGAA
AAAATGTGGTGTGTGACATGTAAAAATGCTCAACCTGGTTTCCAAAGTCTTTCAACGACACC
CTGATCTTCACTAAAAAATTGTAAAGGTTTCAACACGTTGCTTTAATAAATCACTTGCCCTGC

FIGURE 178

MRLSVCLLMVSLALCCYQAHALVCPAVASEITVFLFLSDAAVNLQVAKLNPPPEALAAKLEV KHCTDQISFKKRLSLKKSWWK

FIGURE 179

FIGURE 180

MERVTLALLLLAGLTALEANDPFANKDDPFYYDWKNLQLSGLICGGLLAIAGIAAVLSGKCK YKSSQKQHSPVPEKAIPLITPGSATTC

FIGURE 181

GGAGAAGAGGTTGTGTGGGACAAGCTGCTCCCGACAGAAGG**ATG**TCGCTGCTGAGCCTGCCC $\mathtt{TGGCTGGGCCTCAGACCGGTGGCAATGTCCCCATGGCTACTCCTGCTGCTGGTTGTGGGCTC$ CTGGCTACTCGCCGCATCCTGGCTTGGACCTATGCCTTCTATAACAACTGCCGCCGGCTCC AGTGTTTCCCACAGCCCCCAAAACGGAACTGGTTTTGGGGTCACCTGGGCCTGATCACTCCT ACAGAGGAGGGCTTGAAGGACTCGACCCAGATGTCGGCCACCTATTCCCAGGGCTTTACGGT ATGGCTGGGTCCCATCATCCCCTTCATCGTTTTATGCCACCCTGACACCATCCGGTCTATCA CCAATGCCTCAGCTGCCATTGCACCCAAGGATAATCTCTTCATCAGGTTCCTGAAGCCCTGG CTGGGAGAAGGGATACTGCTGAGTGGCGGTGACAAGTGGAGCCGCCACCGTCGGATGCTGAC GCCCGCCTTCCATTTCAACATCCTGAAGTCCTATATAACGATCTTCAACAAGAGTGCAAACA TCATGCTTGACAAGTGGCAGCACCTGGCCTCAGAGGGCAGCAGTCGTCTGGACATGTTTGAG CACATCAGCCTCATGACCTTGGACAGTCTACAGAAATGCATCTTCAGCTTTGACAGCCATTG TCAGGAGAGGCCCAGTGAATATTTGCCACCATCTTGGAGCTCAGTGCCCTTGTAGAGAAAA GAAGCCAGCATATCCTCCAGCACATGGACTTTCTGTATTACCTCTCCCATGACGGCGGCGC TTCCACAGGGCCTGCCGCCTGGTGCATGACTTCACAGACGCTGTCATCCGGGAGCGGCGTCG CACCCTCCCCACTCAGGGTATTGATGATTTTTTCAAAGACAAAGCCAAGTCCAAGACTTTGG ATTTCATTGATGTGCTTCTGCTGAGCAAGGATGAAGATGGGAAGGCATTGTCAGATGAGGAT ATAAGAGCAGAGGCTGACACCTTCATGTTTGGAGGCCATGACACCACGGCCAGTGGCCTCTC CTGGGTCCTGTACAACCTTGCGAGGCACCCAGAATACCAGGAGCGCTGCCGACAGGAGGTGC AAGAGCTTCTGAAGGACCGCGATCCTAAAGAGATTGAATGGGACGACCTGGCCCAGCTGCCC TTCCTGACCATGTGCGTGAAGGAGAGCCTGAGGTTACATCCCCCAGCTCCCTTCATCTCCCG ATGCTGCACCCAGGACATTGTTCTCCCAGATGGCCGAGTCATCCCCAAAGGCATTACCTGCC TCATCGATATTATAGGGGTCCATCACAACCCAACTGTGTGGCCGGATCCTGAGGTCTACGAC ${\tt CCCTTCCGCTTTGACCCAGAGAACAGCAAGGGGAGGTCACCTCTGGCTTTTATTCCTTTCTC}$ CGCAGGGCCCAGGAACTGCATCGGGCAGGCGTTCGCCATGGCGGAGATGAAAGTGGTCCTGG CGTTGATGCTGCACCTCCGGTTCCTGCCAGACCACTGAGCCCCGCAGGAAGCTGGAA TTGATCATGCGCGCCGAGGGCGGCTTTGGCTGCGGGTGGAGCCCCTGAATGTAGGCTTGCA **GTGA**CTTTCTGACCCATCCACCTGTTTTTTTGCAGATTGTCATGAATAAAACGGTGCTGTCAAA

FIGURE 182

MSLLSLPWLGLRPVAMSPWLLLLLVVGSWLLARILAWTYAFYNNCRRLQCFPQPPKRNWFWG
HLGLITPTEEGLKDSTQMSATYSQGFTVWLGPIIPFIVLCHPDTIRSITNASAAIAPKDNLF
IRFLKPWLGEGILLSGGDKWSRHRRMLTPAFHFNILKSYITIFNKSANIMLDKWQHLASEGS
SRLDMFEHISLMTLDSLQKCIFSFDSHCQERPSEYIATILELSALVEKRSQHILQHMDFLYY
LSHDGRRFHRACRLVHDFTDAVIRERRRTLPTQGIDDFFKDKAKSKTLDFIDVLLLSKDEDG
KALSDEDIRAEADTFMFGGHDTTASGLSWVLYNLARHPEYQERCRQEVQELLKDRDPKEIEW
DDLAQLPFLTMCVKESLRLHPPAPFISRCCTQDIVLPDGRVIPKGITCLIDIIGVHHNPTVW
PDPEVYDPFRFDPENSKGRSPLAFIPFSAGPRNCIGQAFAMAEMKVVLALMLLHFRFLPDHT
EPRRKLELIMRAEGGLWLRVEPLNVGLQ

FIGURE 183

FIGURE 184

MYKLASCCLLFTGFLNPLLSLPLLDSREISFQLSAPHEDARLTPEELERASLLQILPEMLGA ERGDILRKADSSTNIFNPRGNLRKFQDFSGQDPNILLSHLLARIWKPYKKRETPDCFWKYCV

FIGURE 185

FIGURE 186

 ${\tt MPSPGTVCSLLLLGMLWLDLAMAGSSFLSPEHQRVQQRKESKKPPAKLQPRALAGWLRPEDG} \\ {\tt GQAEGAEDELEVRFNAPFDVGIKLSGVQYQQHSQALGKFLQDILWEEAKEAPADKO} \\$

FIGURE 187

CGGCCACAGCTGGCATGCTCTGATCGCCATCCTGCTGTATGTCCTCGTCCAGTACCTC ${\tt GTGAACCCCGGGGTGCTCCGCACGGACCCCAGATGTCAAGAAT} {\color{red} {\bf ATG}} {\tt AACACGTGGCTGCTGT}$ TCCTCCCCTGTTCCCGGTGCAGGTGCAGACCCTGATAGTCGTGATCATCGGGATGCTCGTG $\tt CTCCTGCTGGACTTCTTGGCTTGGTGCACCTGGGCCAGCTGCTCATCTTCCACATCTACCT$ GAGTATGTCCCCCACCCTAAGCCCCCGATCCCCCAAGGCTGGGTGGTCAGAGCTGCTCATC TTACACCTCTACTTGAGTATGTCCCTAACCCTGAGCCCCCACGCCTGGGGCCAGAGTCTTT GTCCCCGTGTGCGCATGTGTTCAGGGTCAGCCTCTCCCAGAAGTGAGATCATGGACAAAAA GGGCAAATCACAGGAAGAAATTAAATCCATGAGGACCCAGCAGGCCCAGCAAGAAGCTGAAC $\verb"TCACGCCGAGACCTGCAGGAGTGCCAGGTGCT" \\ \texttt{TGA} \texttt{AGTAACAAGTTTAAAATGTTCAGA}$ GACAATGGAATGGAATCTATTAGGCAAGAACAGGACATTATGAAATAAGGACAGGTGGACTT AACAACTGAAGCGAGAGCTGTGGTCTTGCTTGGTCTCACAGTGGGCACAGCGGTAGGCGGTC AGTCATGTTGCTGAACGACGGAGGGTAAACTCCCCAGCCCCAAGAAAACCTGTGTTGGAAGT AACAACAACCTCCCTGCTCCTGGCACCAGCCGTTTTGGTCATGGTGGGCCAGCTGCAAAGCG TCTTCCATTCTCTGGGCAGTGGTGGCCCCGAGGCTGTGGCCTCTCAGGGGGTTTCTGTGGAC ACGGGCAGCAGAGTGTGTCCAGGCCAGCCCCCAAGAATGCCCTGCTCCTGACAGCTTGGCCA ACCCCTGGTCAGGGCAGAGGGAGTTGGGTGGGTCAGGCTCTGGGCTCACCTCCATCTCCAGA GCATCCCCTGCCTGCAGTTGTGGCAAGAACGCCCAGCTCAGAATGAACACACCCCACCAAGA GCCTCCTTGTTCATAACCACAGGTTACCCTACAAACCACTGTCCCCACACAACCCTGGGGAT GTTTTAAAACACACACCTCTAACGCATATCTTACAGTCACTGTTGTCTTGCCTGAGGGTTGA ATTTTTTTTAATGAAAGTGCAATGAAAATCACTGGATTAAATCCTACGGACACAGAGCTGAA

FIGURE 188

MNTWLLFLPLFPVQVQTLIVVIIGMLVLLLDFLGLVHLGQLLIFHIYLSMSPTLSPRSPQGW VVRAAHLTPLLEYVPNPEPPTPGARVFVPRVRMCSGSASPRSEIMDKKGKSQEEIKSMRTQQ AQQEAELTPRPAGVVPGA

FIGURE 189

 ${\tt GGAGTGCAGATGGCATCCTTCGGTTCTTCCAGACAAGCTGCAAGACGCTGACC} {\bf \underline{ATG}} {\tt GCCAAG}$ ATGGAGCTCTCGAAGGCCTTCTCTGGCCAGCGGACACTCCTATCTGCCATCCTCAGCATGCT ATCACTCAGCTTCTCCACAACATCCCTGCTCAGCAACTACTGGTTTGTGGGCCACACAGAAGG TGCCCAAGCCCCTGTGCGAGAAAGGTCTGGCAGCCAAGTGCTTTGACATGCCAGTGTCCCTG CCGGTTCTCCTTCCGGAGCTTCCGGAGTGGCATGTGGCTATCCTGTGAGGAAACTGTGGAAG AACCAGGGGAGAGGTGCCGAAGTTTCATTGAACTTACACCACCAGCCAAGAGAGGTGAGAAA GGACTACTGGAATTTGCCACGTTGCAAGGCCCATGTCACCCCACTCTCCGATTTGGAGGGAA GCGGTTGATGGAGAAGGCTTCCCTCCCCTCCCTTGGGGGCTTTGTGGCAAAAATCCTA TGGTTATCCCTGGGAACGCAGATCACCTACATCGGACTTCAATTCATCAGCTTCCTCCTGCT ACTAACAGACTTGCTACTCACTGGGAACCCTGCCTGTGGGCTCAAACTGAGCGCCTTTGCTG CTGTTTCCTCTGTCAGGTCTCCTGGGGATGGTGGCCCACATGATGTATTCACAAGTC TTCCAAGCGACTGTCAACTTGGGTCCAGAAGACTGGAGACCACATGTTTGGAATTATGGCTG GGCCTTCTACATGGCCTGGCTCTCCTTCACCTGCTGCATGGCGTCGGCTGTCACCACCTTCA $ACACGTACACCAGGATGGTGCTGGAGTTCAAGTGCAAGCA{f TAAG}$ TAAGAGCTTCAAGGAAAAC CCGAACTGCCTACCACATCACCATCAGTGTTTCCCTCGGCGGCTGTCAAGTGCAGCCCCCAC CGTGGGTCCTTTGACCAGCTACCACCAGTATCATAATCAGCCCATCCACTCTGTCTCTGAGG AAAGAAGCAGTTAGGTCATCTGTAGAGGAAGAGCAGTGTTAGGAGTTAAGCGGGTTTGGGGA GTAGGCTTGAGCCCTACCTTACACGTCTGCTGATTATCAACATGTGCTTAAGCCAACATCCG TCTCTTGAGCATGGTTTTTAGAGGCTACGAATAAGGCTATGAATAAGGGTTATCTTTAAGTC CTAAGGGATTCCTGGGTGCCACTGCTCTTTTCCTCTACAGCTCCATCTTGTTTCACCCAC CCCACATCTCACACATCCAGAATTCCCTTCTTTACTGATAGTTTCTGTGCCAGGTTCTGGGC TAAACCATGGAGATAAAAAGAAGAGTAAAATACACTTCCCGACCTTAAGGATCTGAAA

FIGURE 190

MAKMELSKAFSGQRTLLSAILSMLSLSFSTTSLLSNYWFVGTQKVPKPLCEKGLAAKCFDMP VSLDGDTNTSTQEVVQYNWETGDDRFSFRSFRSGMWLSCEETVEEPGERCRSFIELTPPAKR GEKGLLEFATLQGPCHPTLRFGGKRLMEKASLPSPPLGLCGKNPMVIPGNADHLHRTSIHQL PPATNRLATHWEPCLWAQTERLCCCFLCPVRSPGDGGPHDVFTSLPSDCQLGSRRLETTCLE LWLGLLHGLALLHLLHGVGCHHLQHVHQDGAGVQVQA

FIGURE 191

AACTGGAAGGAAAGAAAGGTCAGCTTTGGCCCAG**ATG**TGGTTACCCCTTGGTCTCCTG TCTTTATGTCTTTCTCCTATTCTGTCATCTCCCTCACTTAAGTCTCAGGCCTGTCA GCAGCTCCTGTGGACATTGCCATCCCCTCTGGTAGCCTTCAGAGCAAACAGGACAACCTATG TTATGGATGTTTCCACCAACCAGGGTAGTGGCATGGAGCACCGTAACCATCTGTGCTTCTGT GATCTCTATGACAGAGCCACTTCTCCACCTCTGAAATGTTCCCTGCTCTGAAATCTGGCATG GTCTGTTCTCTTATTGTCAACCTCAGCACAACAGGCTGGCGCCAATGGCATTACAGAGAAAG CAATCTGTGTGGGCTAGTGGGCAGATTACCATGCAAGCCCCAGGAGAAATGGAGGAGCTTTGT AGCCACCTCCCTGTCAGCCAGTATTAACATGTCCCCTTCCCCCTGCCCCGCCGTAGATTCAG GACATTCGCCCCTGTGTGCCACCAAACCAGGACTTTCCCCCTTGGCTTGGCATCCCTGGCTCT CTCCTGGTACCCAGCAAGACGTCTGTTCCAGGGCAGTGTAGCATCTTTCAAGCTCCGTTACT ATGGCGATGGCCATGATGTTACAATCCCACTTGCCTGAATAATCAAGTGGGAAGGGGAAGCA GAGGGAAATGGGGCCATGTGAATGCAGCTGCTCTGTTCTCCCTACCCTGAGGAAAAACCAAA TGTTGAAGGGCACAAGAATGTAGCTGGAGAAGATTGATGAAAGTGCAGGTGTGTAAGGAA ATAGAACAGTCTGCTGGGAGTCAGACCTGGAATTCTGATTCCAAACTCTTTATTACTTTGGG AAGTCACTCAGCCTCCCCGTAGCCATCTCCAGGGTGACGGAACCCAGTGTATTACCTGCTGG AACCAAGGAAACTAACAATGTAGGTTACTAGTGAATACCCCAATGGTTTCTCCAATTATGCC CATGCCACCAAAACAATAAAACAAAATTCTCTAACACTGAAA

FIGURE 192

MWLPLGLLSLCLSPLPILSSPSLKSQACQQLLWTLPSPLVAFRANRTTYVMDVSTNQGSGME HRNHLCFCDLYDRATSPPLKCSLL

FIGURE 193

 \mathtt{CCGCC} GGATGTGCCCTTCCAATATACAACAAATACTGGCCCCTCTTTGTTCTATTTTTTTACATCCTTTCACCTATTCC ATACTGCATAGCAAGAAGATTAGTGGATGATACAGATGCTATGAGTAACGCTTGTAAGGAACTTGCCATCTTTC TTACAACGGGCATTGTCGTGTCAGCTTTTGGACTCCCTATTGTATTTGCCAGAGCACATCTGATTGAGTGGGGA GCTTGTGCACTTGTTCTCACAGGAAACACAGTCATCTTTGCAACTATACTAGGCTTTTTCTTGGTCTTTGGAAG ${\tt CAATGACGACTTCAGCTGGCAGCAGTGG} \underline{{\tt TGA}} {\tt AAAGAAATTACTGAACTATTGTCAAATGGACTTCCTGTCATTT}$ GTTGGCCATTCACGCACACAGGAGATGGGGCAGTTAATGCTGAATGGTATAGCAAGCCTCTTGGGGGTATTTTA GGTGCTCCCTTCTCACTTTTATTGTAAGCATACTATTTTCACAGAGACTTGCTGAAGGATTAAAAGGATTTTCT $\tt CTTTTGGAAAAGCTTGACTGATTTCACACTTATCTATAGTATGCTTTTTGTGGTGTCCTGCTGAATTTAAATAT$ TTTGCATTGGTTAGGAATTCAGAATTCCGCCGGCTCTATTACTGGTCAAGTACATCTTTTCTCTTAAAATTATT CAGACATACAGACGGTTGGCATACGTTATAGACTGTATACTCAGTGCAAATATAGCTGCATTTATACCTCAGAG GGGCCAAGTGTTAATGCCCATGCCCTCCGTTAAGGGTTGTTGGTTTACTGGTAGACAGATGTTTTGTGGATTG AAAATTATTTTATGGAATTGCTACAGAGGAGTGCTTTTCTTCTCAATTGTTAGAAGAATTTATGTTAAACTTTA GGGAAGAAATGACATTGAAATTCCAGTTTTTGAATCCTGTTTCTATTTATAAGTGAAATTTGTGATCTCCTATC AACCTTTCATGTTTTACCCTGTTAAAATGGACATACATGGAACCACTACTGATGAGGGACAGTTGTATGTTTGC CAGAGTGCCCCCCCCCCCCAAGGCCTTGCCATGATTAACAAGTAACTTGTTAGTCTTACAGATAATTCATGCA TTAACAGTTTAAGATTTAGACCATGGTAATAGTAGTTCTTATTCTCTAAGGTTATATCATATGTAATTTAAAAG TATTTTTAAGACAAGTTTCCTGTATACCTCTGAACTGTTTTGATTTTGAGTTCATCATGATAGATCTGCTGTTT CCTTATAAAAGGCATTTGTTGTGTGAGTTAATGCAAAGTAGCCAAGTCCAGCTATATAGCAGCTTCAGAAACAT ACCTGACCAAAAAATTCCCAGTAACCAGGCATGATCAATTTATAGTGGTCGTTTACATCTAATAATTATCAGGA AGCCACAGGTTTTATTGCCTAACTTAAGCCATGACTTTTAGATATGAGATGACGGGAAGCAGGACGAAATATCG GCGTGTGGCTGGAGCCTTCCCACTGGAGGCTGAAAGTGGCTTGTGGTATTATAATGTTCAGATTTCAAGAGGAA GGTGCAGGTACACATGAGTTAGAGAGCTGGTGAGACAGTTGGGAACTCTTTGTGCTTGTGATCTACTGGACTTT TTTTTTGCAGGAAGTGCATTCTCTGGTCCTTCCCTATTTTCTGTTCTGGATGTCAGTGCAGTGCACTGCTACTG TTTTATCCACTTGGCCACAGACTTTTTCTAACAGCTGCGTATTATTTCTATATACTAATTGCATTGGCAGCATT GTGTCTTTGACCTTGTATACTAGCTTGACATAGTGCTGTCTCTGATTTCTAGGCTAGTTACTTGAGATATGAAT AGATTTTAAATATCTATTTTAAAAAAAAAA

FIGURE 194

MAGIKALISLSFGGAIGLMFLMLGCALPIYNKYWPLFVLFFYILSPIPYCIARRLVDDTDAM SNACKELAIFLTTGIVVSAFGLPIVFARAHLIEWGACALVLTGNTVIFATILGFFLVFGSND DFSWQQW

FIGURE 195

CTGAAATATATTTTGAGATAGGTTTAGAATGTCA

FIGURE 196

MDFLLLGLCLYWLLRRPSGVVLCLLGACFQMLPAAPSGCPQLCRCEGRLLYCEALNLTEAPH
NLSGLLGLSLRYNSLSELRAGQFTGLMQLTWLYLDHNHICSVQGDAFQKLRRVKELTLSSNQ
ITQLPNTTFRPMPNLRSVDLSYNKLQALAPDLFHGLRKLTTLHMRANAIQFVPVRIFQDCRS
LKFLDIGYNQLKSLARNSFAGLFKLTELHLEHNDLVKVNFAHFPRLISLHSLCLRRNKVAIV
VSSLDWVWNLEKMDLSGNEIEYMEPHVFETVPHLQSLQLDSNRLTYIEPRILNSWKSLTSIT
LAGNLWDCGRNVCALASWLSNFQGRYDGNLQCASPEYAQGEDVLDAVYAFHLCEDGAEPTSG
HLLSAVTNRSDLGPPASSATTLADGGEGQHDGTFEPATVALPGGEHAENAVQIHKVVTGTMA
LIFSFLIVVLVLYVSWKCFPASLRQLRQCFVTQRRKQKQKQTMHQMAAMSAQEYYVDYKPNH
IEGALVIINEYGSCTCHQQPARECEV

FIGURE 197

GTGCAAGGAGCCGAGGCGAGATGGGCCGTCCTGGGCCGGGTCCTGCTGTGGCTGCAGCTCTGC
GCACTGACCCAGGCGGTCTCCAAACTCTGGGTCCCCAACACGGACTTCGACGTCGCAGCCAA
CTGGAGCCAGAACCGGACCCCGTGCGCCGGCGGCGCGCGTTGAGTTCCCGGCGGACAAGATGG
TGTCAGTCCTGGTGCAAGAAGGTCACGCCGTCTCAGACATGCTCCTGCCGCTGGATGGGGAA
CTCGTCCTGGCTTCAGGAGCCGGATTCGGCGTCTCAGACGTGGGCTCGCACCTGGACTGTGG
CGCGGGCGAACCTGCCGTCTTCCGCGACTCTGACCGCTTCTCCTGGCATGACCCGCACCTGT
GGCGCTCTGGGGACGAGGCACCTGGCCTCTTCTTCGTGGACGCCGAGCGCGTGCCCTGCCGC
CACGACGACGTCTTCTTTCCGCCTAGTGCCTCCTTCCGCGTGGGGCTCGGCCCTGGCGCTAG
CCCCGTGCGTGTCCCGCAGCATCTCGGCTCTGGGCCGGACGACGACGACGACCTGG
CTGTTTTCCTGGCGTCCCGCGCGGGCCGCCTACGCTTCCACGGGCCGGGCGCGCCTCTG
GGCCCCGAGGACTGCCGCGCGGCCGCCTACGCTTCCACGGGCCGGAGGCCCAGCCGTG
GGCCCCGAGGACTGCGCGCCCTCCGCCCTCGGCCCTAGCCCTT

FIGURE 198

MGVLGRVLLWLQLCALTQAVSKLWVPNTDFDVAANWSQNRTPCAGGAVEFPADKMVSVLVQE GHAVSDMLLPLDGELVLASGAGFGVSDVGSHLDCGAGEPAVFRDSDRFSWHDPHLWRSGDEA PGLFFVDAERVPCRHDDVFFPPSASFRVGLGPGASPVRVRSISALGRTFTRDEDLAVFLASR AGRLRFHGPGALSVGPEDCADPSGCVCGNAEAQPWICAALLQP

FIGURE 199

FIGURE 200

MGPVKQLKRMFEPTRLIATIMVLLCFALTLCSAFWWHNKGLALIFCILQSLALTWYSLSFIP FARDAVKKCFAVCLA

FIGURE 201

TTGAGCGCAGGTGAGCTCCTGCGCGTTCCGGGGGCGTTCCTCCAGTCACCCTCCCGCCGTTACCCGCGGCGCGC CCGAGGGAGTCTCCTCCAGACCCTCCCTCCCGTTGCTCCAAACTAATACGGACTGAACGGATCGCTGCGAGGGT $\mathtt{AACTGATCAAGTACTTTGAAA}$ \mathtt{ATG} $\mathtt{ACTTCGAAATTTATCTTGGTGTCCTTCATACTTGCTGCACTGAGTCTTTC$ $\overline{\text{AACCACCTTTTCTCTCCAACTAGACCAGCAAAAGGTTCTACTAGTTTCTTTTGATGGATTCCGTTGGGATTACT}$ TATATAAAGTTCCAACGCCCCATTTTCATTATATTATGAAATATGGTGTTCACGTGAAGCAAGTTACTAATGTT TTTATTACAAAAACCTACCCTAACCATTATACTTTGGTAACTGGCCTCTTTGCAGAGAATCATGGGATTGTTGC AAATGATATGTTTGATCCTATTCGGAACAAATCTTTCTCCTTGGATCACATGAATATTTATGATTCCAAGTTTT GGGAAGAAGCGACCAATATGGATCACAAACCAGAGGGCAGGACATACTAGTGGTGCAGCCATGTGGCCCGGA ACAGATGTAAAAATACATAAGCGCTTTCCTACTCATTACATGCCTTACAATGAGTCAGTTTCATTTGAAGATAG AGTTGCCAAAATTGTTGAATGGTTTACGTCAAAAGAGCCCATAAATCTTGGTCTTCTCTATTGGGAAGACCCTG ATGACATGGGCCACCATTTGGGACCTGACAGTCCGCTCATGGGGCCTGTCATTTCAGATATTGACAAGAAGTTA GGATATCTCATACAAATGCTGAAAAAGGCAAAGTTGTGGAACACTCTGAACCTAATCATCACAAGTGATCATGG AATGACGCAGTGCTCTGAGGAAAGGTTAATAGAACTTGACCAGTACCTGGATAAAGACCACTATACCCTGATTG ATCAATCTCCAGTAGCAGCCATCTTGCCAAAAGAAGGTAAATTTGATGAAGTCTATGAAGCACTAACTCACGCT CATCCTAATCTTACTGTTTACAAAAAAGAAGACGTTCCAGAAAGGTGGCATTACAAATACAACAGTCGAATTCA ACCAATCATAGCAGTGGCTGATGAAGGGTGGCACATTTTACAGAATAAGTCAGATGACTTTCTGTTAGGCAACC TCAAAAGAAGCCATGAACTCCACAGATTTGTACCCACTACTATGCCACCTCCTCAATATCACTGCCATGCCACA CAATGGATCATTCTGGAATGTCCAGGATCTGCTCAATTCAGCAATGCCAAGGGTGGTCCCTTATACACAGAGTA CTATACTCCTCCCTGGTAGTGTTAAACCAGCAGAATATGACCAAGAGGGGTCATACCCTTATTTCATAGGGGTC TCTCTTGGCAGCATTATAGTGATTGTATTTTTTGTAATTTTCATTAAGCATTTAATTCACAGTCAAATACCTGC $\tt CTTACAAGATATGCATGCTGAAATAGCTCAACCATTATTACAAGCC{\color{red}{\textbf{TAA}}} TGTTACTTTGAAGTGGATTTGCATA$ ${ t TTGAAGTGGAGATTCCATAATTATGTCAGTGTTTAAAGGTTTCAAA{ t TTCTGGGAAACCAGTTCCAAACATCTGC}$ ATCCTGCTTTATTTGGACTTGGCGCAGATAATGTATATATTTAGCAACTTTGCACTATGTAAAGTACCTTATAT ATTGCACTTTAAATTTCTCTCCTGATGGGTACTTTAATTTGAAATGCACTTTATGGACAGTTATGTCTTATAAC TTGATTGAAAATGACAACTTTTTGCACCCATGTCACAGAATACTTGTTACGCATTGTTCAAACTGAAGGAAATT TCTAATAATCCCGAATAATGAACATAGAAATCTATCTCCATAAATTGAGAGAAGAAGAAGGTGATAAGTGTTGA AAATTAAATGTGATAACCTTTGAACCTTGAATTTTGGAGATGTATTCCCAACAGCAGAATGCAACTGTGGGCAT ATTCGTTCTAAATATTGTTTCTGTCATAAAATTATTGTGATTTCCTGATGAGTCATATTACTGTGATTTCA TAATAATGAAGACACCATGAATATACTTTTCTTCTATATAGTTCAGCAATGGCCTGAATAGAAGCAACCAGGCA AAATCAAATTGGATAAAAAAAAAAAAAAAAAAAAA

FIGURE 202

MTSKFILVSFILAALSLSTTFSLQLDQQKVLLVSFDGFRWDYLYKVPTPHFHYIMKYGVHVK
QVTNVFITKTYPNHYTLVTGLFAENHGIVANDMFDPIRNKSFSLDHMNIYDSKFWEEATPIW
ITNQRAGHTSGAAMWPGTDVKIHKRFPTHYMPYNESVSFEDRVAKIVEWFTSKEPINLGLLY
WEDPDDMGHHLGPDSPLMGPVISDIDKKLGYLIQMLKKAKLWNTLNLIITSDHGMTQCSEER
LIELDQYLDKDHYTLIDQSPVAAILPKEGKFDEVYEALTHAHPNLTVYKKEDVPERWHYKYN
SRIQPIIAVADEGWHILQNKSDDFLLGNHGYDNALADMHPIFLAHGPAFRKNFSKEAMNSTD
LYPLLCHLLNITAMPHNGSFWNVQDLLNSAMPRVVPYTQSTILLPGSVKPAEYDQEGSYPYF
IGVSLGSIIVIVFFVIFIKHLIHSQIPALQDMHAEIAQPLLQA

Signal Peptide:

amino acids 1-22

Transmembrane Domain:

amino acids 429-452

N-glycosylation sites:

amino acids 101-104, 158-161, 292-295, 329-332, 362-365, 369-372, 382-385, 389-392

Somatomedin B Domain:

amino acids 69-85

Sulfatase protein Region:

amino acids 212-241

FIGURE 203

GGATTTTTGTGATCCGCGATTCGCTCCCACGGGCGGACCTTTGTAACTGCGGGAGGCCCAG AGAGAGGCCAAGCCCCTTGCCTTGGGTCACACAGCCAAAGGAGGCAGAGCCAGAACTCACAA $\verb|CCAGATCCAGAGGCAACAGGGACAATGGCCACCTGGGACGAAAAGGCAGTCACCCGCAGGGCC|$ AAGGTGGCTCCCGCTGAGAGGATGAGCAAGTTCTTAAGGCACTTCACGGTCGTGGGAGACGA AGCAGCCACCACCACACCAGTCTCAGGCGAGGAAGGCAGAGCTGCAGCCCCTGACGTTGCC CCTGCCCTGGCCCCGCACCCAGGGCCCCCCTTGACTTCAGGGGCATGTTGAGGAAACTGTT CAGCTCCCACAGGTTTCAGGTCATCATCTGCTTGGTGGTTCTGGATGCCCTCCTGGTGC TTGCTGAGCTCATCCTGGACCTGAAGATCATCCAGCCCGACAAGAATAACTATGCTGCCATG ATTTGTCTTCCGCCTGAGTTCTTTCACCACAAGTTTGAGATCCTGGATGCCCGTCGTGGTGG TGGTCTCATTCATCCTGGACATTGTCCTCCTGTTCCAGGAGCACCAGTTTGAGGCTCTGGGC CTGCTGATTCTGCTCCGGCTGTGGCCGGGTGGCCCGGATCATCAATGGGATTATCATCTCAGT TAAGACACGTTCAGAACGGCAACTCTTAAGGTTAAAACAGATGAATGTACAATTGGCCGCCA AGATTCAACACCTTGAGTTCAGCTGCTCTGAGAAGCCCCTGGAC**TGA**TGAGTTTGCTGTATC AACCTGTAAGGAGAAGCTCTCTCCGGATGGCTATGGGAATGAAAGAATCCGACTTCTACTCT CAGGCTGGCATGTTCACTGGGCTGTTTACGACAGAGAACCTGACAGTCACTGGCCAGTTA TCACTTCAGATTACAAATCACACAGAGCATCTGCCTGTTTTCAATCACAAGAGAACAAAACC AAAATCTATAAAGATATTCTGAAAATATGACAGAATTTGACAAATAAAAGCATAAACGTGTA

FIGURE 204

MATWDEKAVTRRAKVAPAERMSKFLRHFTVVGDDYHAWNINYKKWENEEEEEEEQPPPTPV SGEEGRAAAPDVAPAPGPAPRAPLDFRGMLRKLFSSHRFQVIIICLVVLDALLVLAELILDL KIIQPDKNNYAAMVFHYMSITILVFFMMEIIFKLFVFRLSSFTTSLRSWMPVVVVVSFILDI VLLFQEHQFEALGLLILLRLWRVARIINGIIISVKTRSERQLLRLKQMNVQLAAKIQHLEFS CSEKPLD

FIGURE 205

GCTCTGTCTGTGCCTGTACGTGCCGGTCATCGGGGAAGCCCAGACCGAGTTCCAGTACTTTG AGTCGAAGGGGCTCCCTGCCGAGCTGAAGTCCATTTTCAAGCTCAGTGTCTTCATCCCCTCC CAGGAATTCTCCACCTACCGCCAGTGGAAGCAGAAAATTGTACAAGCTGGAGATAAGGACCT TGATGGGCAGCTAGACTTTGAAGAATTTGTCCATTATCTCCAAGATCATGAGAAGAAGCTGA GGCTGGTGTTTAAGATTTTGGACAAAAAGAATGATGGACGCATTGACGCGCAGGAGATCATG CAGTCCCTGCGGGACTTGGGAGTCAAGATATCTGAACAGCAGCAGAAAAAATTCTCAAGAG CATGGATAAAAACGGCACGATGACCATCGACTGGAACGAGTGGAGAGACTACCACCTCCTCC ACCCCGTGGAAAACATCCCCGAGATCATCCTCTACTGGAAGCATTCCACGATCTTTGATGTG GGTGAGAATCTAACGGTCCCGGATGAGTTCACAGTGGAGGAGAGGCAGACGGGGATGTGGTG GAGACACCTGGTGGCAGGAGGTGGGGCCAGGGGCCGTATCCAGAACCTGCACGGCCCCCCTGG ACAGGCTCAAGGTGCTCATGCAGGTCCATGCCTCCCGCAGCAACAACATGGGCATCGTTGGT GGCTTCACTCAGATGATTCGAGAAGGAGGGGCCAGGTCACTCTGGCGGGGCAATGGCATCAA CGTCCTCAAAATTGCCCCCGAATCAGCCATCAAATTCATGGCCTATGAGCAGATCAAGCGCC TTGTTGGTAGTGACCAGGAGACTCTGAGGATTCACGAGAGGCTTGTGGCAGGGTCCTTGGCA CTGCGCTGCGAGGGTTTCTTTATTTCACTCTTTTCTGAATGTCAAGGCAGTGAGGTGCCTCT

FIGURE 206

MLCLCLYVPVIGEAQTEFQYFESKGLPAELKSIFKLSVFIPSQEFSTYRQWKQKIVQAGDKD
LDGQLDFEEFVHYLQDHEKKLRLVFKILDKKNDGRIDAQEIMQSLRDLGVKISEQQAEKILK
SMDKNGTMTIDWNEWRDYHLLHPVENIPEIILYWKHSTIFDVGENLTVPDEFTVEERQTGMW
WRHLVAGGGAGAVSRTCTAPLDRLKVLMQVHASRSNNMGIVGGFTQMIREGGARSLWRGNGI
NVLKIAPESAIKFMAYEQIKRLVGSDQETLRIHERLVAGSLAGAIAQSSIYPMEVLKTRMAL
RKTGQYSGMLDCARRILAREGVAAFYKGYVPNMLGIIPYAGIDLAVYETLKNAWLQHYAVNS
ADPGVFVLLACGTMSSTCGQLASYPLALVRTRMQAQASIEGAPEVTMSSLFKHILRTEGAFG
LYRGLAPNFMKVIPAVSISYVVYENLKITLGVQSR

Important features:

Signal peptide:

amino acids 1-16

Transmembrane domain:

amino acids 284-304, 339-360, 376-394

Mitochondrial energy transfer proteins signature.

amino acids 206-215, 300-309

N-glycosylation site.

amino acids 129-133, 169-173

Elongation Factor-hand calcium-binding protein.

amino acids 54-73, 85-104, 121-140

FIGURE 207

GGAAGGCAGCGGCAGCTCCACTCAGCCAGTACCCAGATACGCTGGGAACCTTCCCCAGCCAT CAATTGCACTCATCATTGGCTTTGGTATTTCAGGGAGACACTCCATCACAGTCACTACTGTC GCCTCAGCTGGGAACATTGGGGAGGATGGAATCCTGAGCTGCACTTTTGAACCTGACATCAA ACTTTCTGATATCGTGATACAATGGCTGAAGGAAGGTGTTTTAGGCTTGGTCCATGAGTTCA AAGAAGGCAAAGATGAGCTGTCGGAGCAGGATGAAATGTTCAGAGGCCGGACAGCAGTGTTT GCTGATCAAGTGATAGTTGGCAATGCCTCTTTGCGGCTGAAAAACGTGCAACTCACAGATGC TGGCACCTACAAATGTTATATCATCACTTCTAAAGGCAAGGGGAATGCTAACCTTGAGTATA AAACTGGAGCCTTCAGCATGCCGGAAGTGAATGTGGACTATAATGCCAGCTCAGAGACCTTG CGGTGTGAGGCTCCCCGATGGTTCCCCCAGCCCACAGTGGTCTGGGCATCCCAAGTTGACCA GGGAGCCAACTTCTCGGAAGTCTCCAATACCAGCTTTGAGCTGAACTCTGAGAATGTGACCA TGAAGGTTGTCTGTGCTCTACAATGTTACGATCAACAACACATACTCCTGTATGATTGAA AATGACATTGCCAAAGCAACAGGGGATATCAAAGTGACAGAATCGGAGATCAAAAGGCGGAG GGGCACTTCTGCCTCTCAGCCCTTACCTGATGCTAAAA**TAA**TGTGCCTTGGCCACAAAAAAG CATGCAAAGTCATTGTTACAACAGGGATCTACAGAACTATTTCACCACCAGATATGACCTAG TTTTATATTTCTGGGAGGAAATGAATTCATATCTAGAAGTCTGGAGTGAGCAAACAAGAGCA GACATATTAGAAGTTGGGAAAATAATTCATGTGAACTAGACAAGTGTGTTAAGAGTGATAAG GGGGAGTGAGAGGACAGGATAGTGCATGTTCTTTGTCTCTGAATTTTTAGTTATATGTGCTG TAATGTTGCTCTGAGGAAGCCCCTGGAAAGTCTATCCCAACATATCCACATCTTATATTCCA CAAATTAAGCTGTAGTATGTACCCTAAGACGCTGCTAATTGACTGCCACTTCGCAACTCAGG GGCGGCTGCATTTTAGTAATGGGTCAAATGATTCACTTTTTATGATGCTTCCAAAGGTGCCT TGGCTTCTCTCCCAACTGACAAATGCCAAAGTTGAGAAAAATGATCATAATTTTAGCATAA ACAGAGCAGTCGGGGACACCGATTTTATAAATAAACTGAGCACCTTCTTTTTAAACAAAAAA

FIGURE 208

MASLGQILFWSIISIIIILAGAIALIIGFGISGRHSITVTTVASAGNIGEDGILSCTFEPDI KLSDIVIQWLKEGVLGLVHEFKEGKDELSEQDEMFRGRTAVFADQVIVGNASLRLKNVQLTD AGTYKCYIITSKGKGNANLEYKTGAFSMPEVNVDYNASSETLRCEAPRWFPQPTVVWASQVD QGANFSEVSNTSFELNSENVTMKVVSVLYNVTINNTYSCMIENDIAKATGDIKVTESEIKRR SHLQLLNSKASLCVSSFFAISWALLPLSPYLMLK

FIGURE 209

GAATTTGTAGAAGACAGCGGCGTTGCC**ATG**GCGGCGTCTCTGGGGCAGGTGTTGGCTCTGGT GCTGGTGGCCGCTCTGTGGGGTGGCACGCAGCCGCTGCTGAAGCGGGCCTCCGCCGGCCTGC AGCGGGTTCATGAGCCGACCTGGGCCCAGCAGTTGCTACAGGAGATGAAGACCCTCTTCTTG AATACTGAGTACCTGATGCCCTTTCTCCTCAACCAGTGTGGATCCCTTCTCTATTACCTCAC CTTGGCATCGACAGATCTGACCCTGGCTGTGCCCATCTGTAACTCTCTGGCTATCATCTTCA CACTGATTGTTGGGAAGGCCCTTGGAGAAGATATTGGTGGAAAACGTAAGTTAGACTACTGC CTCCCCAGAGTGGGTGAGGACACGGCCTTTTCCCATCCTGCCCTTTCCTCTGCAGCTGTTTT GCTTCCTTGTGGCCATCAGAGTTCCCTTCCCCTGGACAGTCTGGAGAAAGACAGAGGCTGGG GTTTGGGAT**TGA**GACCAGACCCCATCTGAGCCCTTCCTCCAGCCCTGTACCAGCTCCTACT GGCATGGCTGAGCTCAGACCCTCCTGATTTCTGCCTATTATCCCAGGAGCAGTTGCTGGCAT GGTGCTCACCGTGATAGGAATTTCACTCTGCATCACAAGCTCAGTGAGTAAGACCCAGGGGC AACAGTCTACCCTTTGAGTGGGCCGAACCCACTTCCAGCTCTGCTGCCTCCAGGAAGCCCCT GGGCCATGAAGTGCTGGCAGTGAGCGGATGGACCTAGCACTTCCCCTCTCTGGCCTTAGCTT CCTCCTCTTTATGGGGATAACAGCTACCTCATGGATCACAATAAGAGAACAAGAGTGAAAG AGTTTTGTAACCTTCAAGTGCTGTTCAGCTGCGGGGATTTAGCACAGGAGACTCTACGCTCA CCCTCAGCAACCTTTCTGCCCCAGCAGCTCTCTTCCTGCTAACATCTCAGGCTCCCAGCCCA GCCACCATTACTGTGGCCTGATCTGGACTATCATGGTGGCAGGTTCCATGGACTGCAGAACT CCAGCTGCATGGAAAGGGCCAGCTGCAGACTTTGAGCCAGAAATGCAAACGGGAGGCCTCTG GGACTCAGTCAGAGCGCTTTGGCTGAATGAGGGGTGGAACCGAGGGAAGAAGGTGCGTCGGA AAATCCTCACTGCCAGCCCCTCTTAAACAGGTAGAGAGCTGTGAGCCCCAGCCCAGCCTGAC

FIGURE 210

MAASLGQVLALVLVAALWGGTQPLLKRASAGLQRVHEPTWAQQLLQEMKTLFLNTEYLMPFL LNQCGSLLYYLTLASTDLTLAVPICNSLAIIFTLIVGKALGEDIGGKRKLDYCECGTQLCGS RHTCVSSFPEPISPEWVRTRPFPILPFPLQLFCFLVAIRVPFPWTVWRKTEAGVWD

FIGURE 211

CTTCTGTAGGACAGTCACCAGGCCAGATCCAGAAGCCTCTCTAGGCTCCAGCTTTCTCTGTG GAAGATGACAGCAATTATAGCAGGACCCTGCCAGGCTGTCGAAAAGATTCCGCAATAAAACT TTGCCAGTGGGAAGTACCTAGTGAAACGGCCTAAGATGCCACTTCTTCTCATGTCCCAGGCT TGAGGCCCTGTGGTCCCCATCCTTGGGAGAAGTCAGCTCCAGCACCATGAAGGGCATCCTCG TTGCTGGTATCACTGCAGTGCTTGTTGCAGCTGTAGAATCTCTGAGCTGCGTGCAGTGTAAT TCATGGGAAAAATCCTGTGTCAACAGCATTGCCTCTGAATGTCCCTCACATGCCAACACCAG CTGTATCAGCTCCTCAGCCAGCTCCTCTCTAGAGACACCAGTCAGATTATACCAGAATATGT TCTGCTCAGCGGAGAACTGCAGTGAGGAGACACACATTACAGCCTTCACTGTCCACGTGTCT GCTGAAGAACACTTTCATTTTGTAAGCCAGTGCTGCCAAGGAAAGGAATGCAGCAACACCAG CGATGCCCTGGACCCTCCCCTGAAGAACGTGTCCAGCAACGCAGAGTGCCCTGCTTGTTATG AATCTAATGGAACTTCCTGTCGTGGGAAGCCCTGGAAATGCTATGAAGAAGAACAGTGTGTC TTTCTAGTTGCAGAACTTAAGAATGACATTGAGTCTAAGAGTCTCGTGCTGAAAAGGCTGTTC CAACGTCAGTAACGCCACCTGTCAGTTCCTGTCTGGTGAAAACAAGACTCTTGGAGGAGTCA TCCCACAACGTGGGCTCCAAAGCTTCCCTCTACCTCTTGGCCCTTGCCAGCCTCCTTCTTCG $\tt GGGACTGCTGCCC{\color{blue}{\textbf{TGA}}} GGTCCTGGGGGCTGCACTTTGCCCAGCACCCCATTTCTGCTTCTCTG$ AGGTCCAGAGCACCCCTGCGGTGCTGACACCCTCTTTCCCTGCTCTGCCCCGTTTAACTGC ${\tt CCAGTAAGTGGGAGTCACAGGTCTCCAGGCAATGCCGACAGCTGCCTTGTTCTTCATTATTA}$

FIGURE 212

MKGILVAGITAVLVAAVESLSCVQCNSWEKSCVNSIASECPSHANTSCISSSASSSLETPVR LYQNMFCSAENCSEETHITAFTVHVSAEEHFHFVSQCCQGKECSNTSDALDPPLKNVSSNAE CPACYESNGTSCRGKPWKCYEEEQCVFLVAELKNDIESKSLVLKGCSNVSNATCQFLSGENK TLGGVIFRKFECANVNSLTPTSAPTTSHNVGSKASLYLLALASLLLRGLLP

FIGURE 213

GGGCTTGCCTCACTGGCCACCCTCCCAACCCCAAGAGCCCAGCCCCATGGTCCCCGCCGCCG GCGCGCTGCTGTGGGTCCTGCTGAATCTGGGTCCCCGGGCGGCGGGGGCCCAAGGCCTG ACCCAGACTCCGACCGAAATGCAGCGGGTCAGTTTACGCTTTGGGGGCCCCATGACCCGCAG CTACCGGAGCACCGCCCGGACTGGTCTTCCCCGGAAGACAAGGATAATCCTAGAGGACGAGA ATGATGCCATGGCCGACGCCGACCGCCTGGCTGGACCAGCGCTGCCGAGCTCTTGGCCGCC ACGGTGTCCACCGGCTTTAGCCGGTCGTCCGCCATTAACGAGGAGGATGGGTCTTCAGAAGA GGGGGTTGTGATTAATGCCGGAAAGGATAGCACCAGCAGAGAGCTTCCCAGTGCGACTCCCA ATACAGCGGGGAGTTCCAGCACGAGGTTTATAGCCAATAGTCAGGAGCCTGAAATCAGGCTG ACTTCAAGCCTGCCGCGCTCCCCCGGGAGGTCTACTGAGGACCTGCCAGGCTCGCAGGCCAC CCTGAGCCAGTGGTCCACACCTGGGTCTACCCCGAGCCGGTGGCCGTCACCCTCACCCACAG CCATGCCATCTCCTGAGGATCTGCGGCTGGTGCTGATGCCCTGGGGCCCGTGGCACTGCCAC TGCAAGTCGGGCACCATGAGCCGGAGCCGGTCTGGGAAGCTGCACGGCCTTTCCGGGCGCCT TCGAGTTGGGGCGCTGAGCCAGCTCCGCACGGAGCACAAGCCTTGCACCTATCAACAATGTC CCTGCAACCGACTTCGGGAAGAGTGCCCCCTGGACACAAGTCTCTGTACTGACACCAACTGT GCCTCTCAGAGCACCACCAGTACCAGGACCACCACTACCCCCTTCCCCACCATCCACCTCAG AAGCAGTCCCAGCCTGCCACCCGCCAGCCCTGCCCAGCCCTGGCTTTTTGGAAACGGGTCA GGATTGGCCTGGAGGATATTTGGAATAGCCTCTCTTCAGTGTTCACAGAGATGCAACCAATA GACAGAACCAGAGGTAATGGCCACTTCATCCACATGAGGAGATGTCAGTATCTCAACCTCT CTTGCCCTTTCAATCCTAGCACCCACTAGATATTTTTAGTACAGAAAAACAAAACTGGAAAA CACAA

FIGURE 214

MVPAAGALLWVLLLNLGPRAAGAQGLTQTPTEMQRVSLRFGGPMTRSYRSTARTGLPRKTRI
ILEDENDAMADADRLAGPAAAELLAATVSTGFSRSSAINEEDGSSEEGVVINAGKDSTSREL
PSATPNTAGSSSTRFIANSQEPEIRLTSSLPRSPGRSTEDLPGSQATLSQWSTPGSTPSRWP
SPSPTAMPSPEDLRLVLMPWGPWHCHCKSGTMSRSRSGKLHGLSGRLRVGALSQLRTEHKPC
TYQQCPCNRLREECPLDTSLCTDTNCASQSTTSTRTTTTPFPTIHLRSSPSLPPASPCPALA
FWKRVRIGLEDIWNSLSSVFTEMQPIDRNQR

FIGURE 215

GCGCCTGGCGAGCGCTCCCAGGGCGACCGTGAGCCGGTGTACCGCGACTGCGTACTGCAGTGCGAAGAGCA GTCGGGACGACTGTAAGTATGAGTGTATGTGGGTCACCGTTGGGCTCTACCTCCAGGAAGGTCACAAAGTGCCT ${\tt CAGTTCCATGGCAAGTGGCCCTTCTCCCGGTTCCTGTTCTTCAAGAGCCGGCATCGGCCGTGGCCTCGTTTCT}$ CAATGGCCTGGCCAGCCTGGTGATGCTCTGCCGCTACCGCACCTTCGTGCCAGCCTCCTCCCCCATGTACCACA $\tt CCTGTGTGGCCTTGGCTGGTTCCCTCAATGCATGGTTCTGGTCCACAGTCTTCCACACCAGGGACACTGAC$ $\tt CTCACAGAGAAAATGGACTACTTCTGTGCCTCCACTGTCATCCTACACTCAATCTACCTGTGCTGCGTCAGGAC$ GTGTGGTGGCTGGCCTGTGGAACCAGCGGCGGCTGCCTCACGTGCGCAAGTGCGTGGTGGTGGTCTT GCTGCTGCAGGGGCTGTCCCTGCTCGAGCTGCTTGACTTCCCACCGCTCTTCTGGGTCCTGGATGCCCATGCCA TCTGGCACATCAGCACCATCCCTGTCCACGTCCTCTTTTTCAGCTTTCTGGAAGATGACAGCCTGTACCTGCTG ${\tt AAGGAATCAGAGGACAAGTTCAAGCTGGAC} {\tt TGA} {\tt AGACCTTGGAGCGAGTCTGCCCCAGTGGGGATCCTGCCCCC}$ GCCCTGCTGGCCTCCCTTCTCCCCTCAACCCTTGAGATGATTTTCTCTTTTCAACTTCTTGAACTTGGACATGA AGGATGTGGGCCCAGAATCATGTGGCCAGCCCACCCCTGTTGGCCCTCACCAGCCTTGGAGTCTGTTCTAGGG AAGGCCTCCCAGCATCTGGGACTCGAGAGTGGGCAGCCCCTCTACCTCCTGGAGCTGAACTGGGGTGGAACTGA GTGTGTTCTTAGCTCTACCGGGAGGACAGCTGCCTGTTTCCTCCCCACCAGCCTCCCCCACATCCCCAGCTG $\verb|CCTGGCTGGGTCCTGAAGCCCTCTGTCTACCTGGGAGACCAGGGACCACAGGCCTTAGGGATACAGGGGGTCCC||$ GGTTCACGGCGATTCTCCCCATGGGATCTTGAGGGACCAAGCTGCTGGGATTGGGAAGGAGTTTCACCCTGACC GTTGCCCTAGCCAGGTTCCCAGGAGGCCTCACCATACTCCCTTTCAGGGCCAGGGCTCCAGCAAGCCCAGGGCA AGGATCCTGTGCTGCTGTTGAGAGCCTGCCACCGTGTGTCGGGAGTGTGGGCCAGGCTGAGTGCATAGG TGACAGGGCCGTGAGCATGGGCCTGGGTGTGTGTGAGCTCAGGCCTAGGTGCGCAGTGTGGAGACGGGTGTTGT TGCGCGTGCTGGTGGGCATGTGAGATGACTGCCGGTGAATGTGTCCACAGTTGAGAGGTTGGAGCAGGAT GAGGGAATCCTGTCACCATCAATAATCACTTGTGGAGCGCCAGCTCTGCCCAAGACGCCACCTGGGCGGACAGC ${\tt CAGGAGCTCTCCATGGCCAGGCTGCCTGTGTGCATGTTCCCTGTCTGGTGCCCCTTTGCCCGCCTCCTGCAAAC}$ $\verb|CTCACAGGGTCCCCACACACAGTGCCCTCCAGAAGCAGCCCCTCGGAGGCAGAGGAAAATGGGGATGGC||$ ACACTATGCCTGTGCCCTGGTAAAGGTGACCCCTGCCATTTACCAGCAGCCCTGGCATGTTCCTGCCCCACAGG AATAGAATGGAGGGAGCTCCAGAAACTTTCCATCCCAAAGGCAGTCTCCGTGGTTGAAGCAGACTGGATTTTTG CTCTGCCCCTGACCCCTTGTCCCTCTTTGAGGGAGGGGAGCTATGCTAGGACTCCAACCTCAGGGACTCGGGTG GCCTGCGCTAGCTTCTTTTGATACTGAAAACTTTTAAGGTGGGAGGGTGGCAAGGGATGTGCTTAATAAATCAA TTCCAAGCCTCAAAAAAAAAAAAAAAAAA

FIGURE 216

MAGLAARLVLLAGAAALASGSQGDREPVYRDCVLQCEEQNCSGGALNHFRSRQPIYMSLAGW
TCRDDCKYECMWVTVGLYLQEGHKVPQFHGKWPFSRFLFFQEPASAVASFLNGLASLVMLCR
YRTFVPASSPMYHTCVAFAWVSLNAWFWSTVFHTRDTDLTEKMDYFCASTVILHSIYLCCVR
TVGLQHPAVVSAFRALLLLMLTVHVSYLSLIRFDYGYNLVANVAIGLVNVVWWLAWCLWNQR
RLPHVRKCVVVVLLLQGLSLLELLDFPPLFWVLDAHAIWHISTIPVHVLFFSFLEDDSLYLL
KESEDKFKLD

Important features:

Signal peptide:

amino acids 1-20

Transmembrane domains:

amino acids 105-123, 138-156, 169-185, 193-209, 221-240, 256-272

N-glycosylation site.

amino acids 40-44

N-myristoylation site.

amino acids 43-49

CUB domain proteins profile.

amino acids 285-302

Amiloride-sensitive sodium channels proteins.

amino acids 162-186

FIGURE 217

GGCCGCCTGGAATTGTGGGAGTTGTGTCTGCCACTCGGCTGCCGGAGGCCGAAGGTCCGTGA CTATGCTCCCCAGAGCCTGCCTTCATCTAGGATGCTCCTCTGGGCATGCTGCTTGGGCTG CTGATGGCCGCCTGCTTCACCTTCTGCCTCAGTCATCAGAACCTGAAGGAGTTTGCCCTGAC TGGATGCCGAAGTCCTGGAGGTGTTCCACCCGACGCATGAGTGGCAGGCCCTTCAGCCAGGG ACTCCAATATGAGGACAAGTTCCGAAATAATTTGAAAGGCAAAAGGCTGGATATCAACACCA ACACCTACACATCTCAGGATCTCAAGAGTGCACTGGCAAAATTCAAGGAGGGGGCAGAGATG GAGAGTTCAAAGGAAGACAAGGCAAGGCAGGCTGAGGTAAAGCGGCTCTTCCGCCCCATTGA GGAACTGAAGAAAGACTTTGATGAGCTGAATGTTGTCATTGAGACTGACATGCAGATCATGG TACGGCTGATCAACAAGTTCAATAGTTCCAGCTCCAGTTTGGAAGAAGATTGCTGCGCTC TTTGATCTTGAATATTATGTCCATCAGATGGACAATGCGCAGGACCTGCTTTCCTTTGGTGG TCTTCAAGTGGTGATCAATGGGCTGAACAGCACAGAGCCCCTCGTGAAGGAGTATGCTGCGT TTGTGCTGGGCGCTGCCTTTTCCAGCAACCCCAAGGTCCAGGTGGAGGCCATCGAAGGGGGA GCCCTGCAGAAGCTGCTCGTCATCCTGGCCACGGAGCAGCCGCTCACTGCAAAGAAGAAGAT CCTGTTTGCACTGTGCTCCCTGCTGCGCCACTTCCCCTATGCCCAGCGGCAGTTCCTGAAGC TCGGGGGGCTGCAGGTCCTGAGGACCCTGGTGCAGGAGAGGGCACGGAGGTGCTCGCCGTG CGCGTGGTCACACTGCTCTACGACCTGGTCACGGAGAAGATGTTCGCCGAGGAGGAGGCTGA GCTGACCCAGGAGATGTCCCCAGAGAAGCTGCAGCAGTATCGCCAGGTACACCTCCTGCCAG GCCTGTGGGAACAGGGCTGGTGCGAGATCACGGCCCACCTCCTGGCGCTGCCCGAGCATGAT GCCCGTGAGAAGGTGCTGCAGACACTGGGCGTCCTCCTGACCACCTGCCGGGACCGCTACCG TCAGGACCCCCAGCTCGGCAGGACACTGGCCAGCCTGCAGGCTGAGTACCAGGTGCTGGCCA GCCTGGAGCTGCAGGATGGTGAGGACGAGGGCTACTTCCAGGAGCTGCTGGGCTCTGTCAAC AGCTTGCTGAAGGAGCTGAGATGAGCCCCACACCAGGACTGGACTGGATGCCGCTAGTGA GGCTGAGGGTGCCAGCGTGGGTGGGCTTCTCAGGCAGGAGGACATCTTGGCAGTGCTGGCT

FIGURE 218

MAPQSLPSSRMAPLGMLLGLLMAACFTFCLSHQNLKEFALTNPEKSSTKETERKETKAEEEL
DAEVLEVFHPTHEWQALQPGQAVPAGSHVRLNLQTGEREAKLQYEDKFRNNLKGKRLDINTN
TYTSQDLKSALAKFKEGAEMESSKEDKARQAEVKRLFRPIEELKKDFDELNVVIETDMQIMV
RLINKFNSSSSSLEEKIAALFDLEYYVHQMDNAQDLLSFGGLQVVINGLNSTEPLVKEYAAF
VLGAAFSSNPKVQVEAIEGGALQKLLVILATEQPLTAKKKVLFALCSLLRHFPYAQRQFLKL
GGLQVLRTLVQEKGTEVLAVRVVTLLYDLVTEKMFAEEEAELTQEMSPEKLQQYRQVHLLPG
LWEQGWCEITAHLLALPEHDAREKVLQTLGVLLTTCRDRYRQDPQLGRTLASLQAEYQVLAS
LELQDGEDEGYFQELLGSVNSLLKELR

Important features:

Signal peptide:

amino acids 1-29

Hypothetical YJL126w/YLR351c/yhcX family protein.

amino acids 364-373

N-glycosylation site.

amino acids 193-197, 236-240

N-myristoylation site.

amino acids 15-21, 19-25, 234-240, 251-257, 402-408, 451-457

Homologous region SLS1 protein.

amino acids 68-340

FIGURE 219

TTCGGCTTCCGTAGAGGAAGTGGCGCGGACCTTCATTTGGGGTTTCGGTTCCCCCCCTTCCC CTTCCCCGGGGTCTGGGGGTGACATTGCACCGCGCCCCTCGTGGGGTCGCGTTGCCACCCCA CGCGGACTCCCAGCTGGCGCCCCCCCCATTTGCCTGTCCTGGTCAGGCCCCCACCCCCC $\tt TTCCCACCTGACCAGCC \textbf{ATG} \tt GGGGGCTGCGGTTTTTCGGCTGCACTTTCGTCGCGTTCGGCTTCGGCTTCGGCTTCGGCTTCGGCTTCGGCTTCGGCTGCACTTTCGTCGGCTGCGTTCGGCTTTCGGCTGCACTTTCGTCGGCTGCGCTTCGGCTGCACTTTCGGCGTGTTTTTCGGCTGCACTTTCGTCGGCTGTTCGGCTGCACTTTCGGCCTGCACTTCGGCGTGTTCGGCGTGTTTTTTCGGCTGCACTTTCGTCGGCTGTTCGGCTGCACTTTCGGCGTGTTCGGCGTGTTTTTTCGGCTGCACTTTCGTCGGCTGTCGGCTGCACTTTCGGCGTGTTCGGCGCTGCACTTTCGGCGCTGCACTTTCGGCGTGTTCGGCGCTGCACTTTCGGCGTGTTCGGCGTGTTTTTTCGGCCTGCACTTTCGGCGGTGTTCGGCGTGTTTTTTCGGCCTGCACTTTCGGCGTGTTCGGCGTGTTTTTTCGGCCTGCACTTTCGGCGTGTTCGGCGTGTTTTTTCGGCCTGCACTTTCGGCGTGTTCGGCGTGTTTTTTCGGCTGCACTTTCGGCGTGTTCGGCGTGTTTTTTCGGCCTGCACTTTCGGCGTGTTCGGCGTGTTTTTTCGGCTGCACTTTCGGCTGCGTGTTCGGCGTGTTTTTTCGGCTGCACTTTCGGCTGCGTGTTCGGCTGCACTTTCGGCTGCACTTTCGGCTGCACTTTCGGCCGGTGTTTTTTCGGCTGCACTTTCGGCCTGCACTTTCGGCCTGCACTTCGGCTGCACTTTCGGCTGCACTTTCGGCCTGCACTTTCGGCCTGCACTTTCGGCCTGCACTTTCGGCCTGCACTTTCACTT$ CCGGCCTTCGCGCTTTTCTTGATCACTGTGGCTGGGGACCCGCTTCGCGTTATCATCCTGGT CGCAGGGGCATTTTTCTGGCTGGTCTCCTGCTCCTGGCCTCTGTGGTCTGGTTCATCTTGG TCCATGTGACCGACCGGTCAGATGCCCGGCTCCAGTACGCCTCCTGATTTTTGGTGCTGCT GTCTCTGTCCTTCTACAGGAGGTGTTCCGCTTTGCCTACTACAAGCTGCTTAAGAAGGCAGA TGAAGGGTTAGCATCGCTGAGTGAGGACGGAAGATCACCCATCTCCATCCGCCAGATGGCCT ATGTTTCTGGTCTCCTTCGGTATCATCAGTGGTGTCTTCTCTGTTATCAATATTTTGGCT GATGCACTTGGGCCAGGTGTGGTTGGGATCCATGGAGACTCACCCTATTACTTCCTGACTTC AGCCTTTCTGACAGCAGCCATTATCCTGCTCCATACCTTTTGGGGAGTTGTGTTCTTTGATG CCTGTGAGAGGGGGTACTGGGCTTTGGGCCTGGTGGTTGGGAGTCACCTACTGACATCG GGACTGACATTCCTGAACCCCTGGTATGAGGCCAGCCTGCTGCCCATCTATGCAGTCACTGT TTCCATGGGGCTCTGGGCCTTCATCACAGCTGGAGGGTCCCTCCGAAGTATTCAGCGCAGCC TCTTGTGTAAGGAC**TGA**CTACCTGGACTGATCGCCTGACAGATCCCACCTGCCTGTCCACTG CCCATGACTGAGCCCAGCCCGGGTCCATTGCCCACATTCTCTGTCTCCTTCTCGTC GGTCTACCCCACTACCTCCAGGGTTTTGCTTTGTCCTTTTGTGACCGTTAGTCTCTAAGCTT TACCAGGAGCAGCCTGGGTTCAGCCAGTCAGTGACTGGTGTGTTTGAATCTGCACTTATCCC CACCACCTGGGGACCCCCTTGTTGTGTCCAGGACTCCCCCTGTGTCAGTGCTCTGCTCTCAC CCTGCCCAAGACTCACCTCCCTTCCCCTCTGCAGGCCGACGCAGGAGGACAGTCGGGTGAT GGTGTATTCTGCCCTGCGCATCCCACCCGAGGACTGAGGGAACCTAGGGGGGACCCCTGGGC CTGGGGTGCCCTCTGATGTCCTCGCCCTGTATTTCTCCATCTCCAGTTCTGGACAGTGCAG GTTGCCAAGAAAAGGGACCTAGTTTAGCCATTGCCCTGGAGATGAAATTAATGGAGGCTCAA GGATAGATGAGCTCTGAGTTTCTCAGTACTCCCTCAAGACTGGACATCTTGGTCTTTTTCTC GAGGTGGGGGGGGGGGGGGTATATTGGAACTCTTCTAACCTCCTTGGGCTATATTTTCTC TCCTCGAGTTGCTCCTCATGGCTGGGCTCATTTCGGTCCCTTTCTCCTTGGTCCCAGACCTT GGGGGAAAGGAAGTGCATGTTTGGGAACTGGCATTACTGGAACTAATGGTTTTAACCT CCTTAACCACCAGCATCCCTCTCCCCAAGGTGAAGTGGAGGGTGCTGTGGTGAGCTGGC CACTCCAGAGCTGCAGTGCCACTGGAGGAGTCAGACTACCATGACATCGTAGGGAAGGAGGG ATCATTTTCTGCTGAGGGTGGAGTGTCCCATCCTTTTAATCAAGGTGATTGTGATTTTGACT

FIGURE 220

MGAAVFFGCTFVAFGPAFALFLITVAGDPLRVIILVAGAFFWLVSLLLASVVWFILVHVTDR SDARLQYGLLIFGAAVSVLLQEVFRFAYYKLLKKADEGLASLSEDGRSPISIRQMAYVSGLS FGIISGVFSVINILADALGPGVVGIHGDSPYYFLTSAFLTAAIILLHTFWGVVFFDACERRR YWALGLVVGSHLLTSGLTFLNPWYEASLLPIYAVTVSMGLWAFITAGGSLRSIQRSLLCKD

FIGURE 221

FIGURE 222

GACCGACCGTTCAGATGCCCGGTTCCAGTACGGCTTCCTGATTTTTGGTGCTGCTGTNTCTG
TCCTTCTACAGGAGGTGTTCCGCTTTGCCTANTACAAGCTGCTTAAGAAGGCAGATGAGGGG
TTAGCATNGCTGAGTGAGGACGGAAGATCACCCATTTCCATCCGCCAGATGGCCTATGTTTN
TGGTNTTTCCTTCGGTATCATCAGTGGTGTTTTNTCTGTTATCAATATTTTGGNTGATGCAN
TTGGGCCAGGTGTGGTTGGGATCCATGGAGANTCACCCTATTAATTCCTGAATTCAGCCTTT
NTGACAGCAGCCATTATCCTGNTCCATACCTTTTGGGGAGTTGTGTTTTTTGATGCCTGTGA
GAGGAG

FIGURE 223

NGTTGGAGAAGTGGCGCGGACNTTCATTTGGGGTTTCCCCCCTTTCCCTTTCCCCG
GGGTCTGGGGTGACATTGCACGGGCCCCTCGTGGGGTCGCGTTGCCACCCCACGCGGACTCC
CCAGNTGGNGCGCCCTTCCCATTTGCCTGTCCTGGTCAGGCCCCCACCCCCCTTCCCACNTG
ACCAGCCATGGGGGGCTGCGGTGTTTTTCGGCTGCACTTTCGTCGCGTTCGGCCCGGCCTTCG
CGCTTTTCTTGATCACTGTGGCTGGGGACCCGCTTCGCGTTATCATCCTGGTCGCAGGGGCA
TTTTTCTGGCTGGTCTCCCTGCTCCTGGCCTCTGTGGTTCATCTTTGGTCCATGTGAC
CGACCGGTCAGATGCCCGGCTCCAGTACGGCCTCCTGATTTTTGGTGCTGCTGTCTCTGTCC
TTCTACAGGAGGTGTTCCGCTTTGCCTACTACAAGCTGCTTAAGAAGGCAGATGAGGGGTTA
GCATCGCTGAGTGAGGACGGAAGATCACCCATCTCCATCCGCCAGATGGCCTATGTTTCTGG
TCTCTCCTTCGGTATCATCAGTGGTGTCTTCTCTGTTATCAATATTTTGGCTGATGCACTTG
GGCCAGGTGTGGTTTGGGATCCATGGAGACTCACCC

FIGURE 224

GTAAAAGAAAGTGGCCGGACCTTCATTGGGGTTTCGCTTCCCCCTTTCCCNTTCCCCGGGG
TCTGGGGGTGACATTGCACCGCGCCCNTCGTGGGGTCGCGTTGCCACCCCACGCGGACTCCC
CAGNTGGCGCGCCCCCCCCATTTGCCTGTCCTGGTCAGGCCCCCACCCCCCTTCCCACCTGA
CCAGCCATGGGGGCTGCGGTGTTTTTCGGGCTGCACTTTCGTCGCGTTCGGGCCCGGCCTTC
GCGCTTTTCTTGATCACTGTGGCTGGGGACCCGCTTCGCGTTATCATCCTGGTCGCAGGGGC
ATTTTTCTGGCTGGTCTCCCTGCTCCTGGCCTCTGTGGTCATCTTTGGTCCATGTGA
CCGACCGGTCAGATGCCCGGCTCCAGTACGGCCTCCTGATTTTTGGTGCTGCTTCTCTGTC
CTTCTACAGGAGGTGTTCCGCTTTGCCTACTACAAGCTGCTTAAGAAGGCAGATGAGGGGTT
AGCATCGCTGAGTGAGGACGGAAGATCACCCATCTCCATCCGCCAGATGGCCTATGTTTCTG
GTCTCTCCTTCGGTATCATCAGTGGTGTCTTCTCTGTTATCAATATTTTGGCTGATGCACTT
GGGCCAGGTGTGGTTGGGATCCATGGAGAC

FIGURE 225

GCCCCAGGGAGCAGTGGGTGGTTATAACTCAGGCCCGGTGCCCAGAGCCCAGGAGGAGGAGGCAG TGGCCAGGAAGGCACAGGCCTGAGAAGTCTGCGGCTGAGCTGGGAGCAAATCCCCCACCCCC TGTCTGTGCGTCCTGCACCCACATCTTTCTCTGTCCCCTCCTTGCCCTGTCTGGAGGCTGCT AGACTCCTATCTTCTGAATTCTATAGTGCCTGGGTCTCAGCGCAGTGCCGATGGTGGCCCGT $\verb|CCTTGTGGTTCCTCTACCTGGGGAAATAAGGTGCAGCGGCC| \textbf{ATG} \\ \texttt{GCTACAGCAAGACCCC} \\$ CCTGGATGTGGGTGCTCTGTGCTCTGATCACAGCCTTGCTTCTGGGGGGTCACAGAGCATGTT CTCGCCAACAATGATGTTTCCTGTGACCACCCCTCTAACACCGTGCCCTCTGGGAGCAACCA GGACCTGGGAGCTGGGGCCGGGAAGACGCCCGGTCGGATGACAGCAGCAGCCGCATCATCA ATGGATCCGACTGCGATATGCACACCCAGCCGTGGCAGGCCGCGCTGTTGCTAAGGCCCAAC CAGCTCTACTGCGGGGCGGTGTTGGTGCATCCACAGTGGCTGCTCACGGCCGCCCACTGCAG GAAGAAGTTTTCAGAGTCCGTCTCGGCCACTACTCCCTGTCACCAGTTTATGAATCTGGGC AGCAGATGTTCCAGGGGGTCAAATCCATCCCCACCCTGGCTACTCCCACCCTGGCCACTCT AACGACCTCATGCTCATCAAACTGAACAGAAGAATTCGTCCCACTAAAGATGTCAGACCCAT CCAAGAGCCCCCAAGTGCACTTCCCTAAGGTCCTCCAGTGCTTGAATATCAGCGTGCTAAGT CAGAAAAGGTGCGAGGATGCTTACCCGAGACAGATAGATGACACCATGTTCTGCGCCGGTGA CAAAGCAGGTAGAGACTCCTGCCAGGGTGATTCTGGGGGGCCTGTGGTCTGCAATGGCTCCC TGCAGGGACTCGTGTCCTGGGGAGATTACCCTTGTGCCCGGCCCAACAGACCGGGTGTCTAC ACGAACCTCTGCAAGTTCACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCC**TGA**GTCAT CCCAGGACTCAGCACCGGCATCCCCACCTGCTGCAGGGACAGCCCTGACACTCCTTTCAG ACCCTCATTCCTCCCAGAGATGTTGAGAATGTTCATCTCTCCAGCCCCTGACCCCATGTCT CCTGGACTCAGGGTCTGCTTCCCCCACATTGGGCTGACCGTGTCTCTAGTTGAACCCTGG GAACAATTTCCAAAACTGTCCAGGGCGGGGGTTGCGTCTCAATCTCCCTGGGGCACTTTCAT CCTCAAGCTCAGGGCCCATCCCTTCTCTGCAGCTCTGACCCAAATTTAGTCCCAGAAATAAA CTGAGAAGTGGAAAAAAAA

FIGURE 226

MATARPPWMWVLCALITALLLGVTEHVLANNDVSCDHPSNTVPSGSNQDLGAGAGEDARSDD SSSRIINGSDCDMHTQPWQAALLLRPNQLYCGAVLVHPQWLLTAAHCRKKVFRVRLGHYSLS PVYESGQQMFQGVKSIPHPGYSHPGHSNDLMLIKLNRRIRPTKDVRPINVSSHCPSAGTKCL VSGWGTTKSPQVHFPKVLQCLNISVLSQKRCEDAYPRQIDDTMFCAGDKAGRDSCQGDSGGP VVCNGSLQGLVSWGDYPCARPNRPGVYTNLCKFTKWIQETIQANS

FIGURE 227

ATGGTCAACGACCGGTGGAAGACCATGGGCGCGCCCCAACTTGAGGACCGGCCGCGCA $\overline{ ext{CAA}}$ GCCGCAGCCGAGCTGCGGCTACGTGCTGTGCACCGTGCTGCTGGCCCTGGCTGTGC TGCTGGCTGTAGCTGTCACCGGTGCCGTGCTCTTCCTGAACCACGCCCACGCGCCGGGCACG ACGGCTCCGTGAACTTCTTCCGGGGCTGGGACGCGTACCGAGACGGCTTTGGCAGGCTCACC GCCTAGCAGCCTCTCCTCGGGCAGGAGGGGGGGGGGGTGGCTTCCTCCAAAGGACACCCGATGGCA GGTGCCTAGGGGGTGTGGGGTTCCGTTCTCCCTTCCCCACTGAAGTTTGTGCTTAAAA AACAATAAATTTGACTTGGCACCACTGGGGGGTTGGTGGGGAGAGGCCGTGTGACCTGGCTCTC TGTCCCAGTGCCACCAGGTCATCCACATGCGCAG

FIGURE 228

MVNDRWKTMGGAAQLEDRPRDKPQRPSCGYVLCTVLLALAVLLAVAVTGAVLFLNHAHAPGT
APPPVVSTGAASANSALVTVERADSSHLSILIDPRCPDLTDSFARLESAQASVLQALTEHQA
QPRLVGDQEQELLDTLADQLPRLLARASELQTECMGLRKGHGTLGQGLSALQSEQGRLIQLL
SESQGHMAHLVNSVSDILDALQRDRGLGRPRNKADLQRAPARGTRPRGCATGSRPRDCLDVL
LSGQQDDGVYSVFPTHYPAGFQVYCDMRTDGGGWTVFQRREDGSVNFFRGWDAYRDGFGRLT
GEHWLGLKRIHALTTQAAYELHVDLEDFENGTAYARYGSFGVGLFSVDPEEDGYPLTVADYS
GTAGDSLLKHSGMRFTTKDRDSDHSENNCAAFYRGAWWYRNCHTSNLNGQYLRGAHASYADG
VEWSSWTGWQYSLKFSEMKIRPVREDR

FIGURE 229

GCAGTCAGAGACTTCCCCTGCCCCTCGCTGGGAAAGAACATTAGGAATGCCTTTTAGTGCCT TGCTTCCTGAACTAGCTCACAGTAGCCCGGCGGCCCAGGGCAATCCGACCACATTTCACTCT CACCGCTGTAGGAATCCAGATGCAGGCCAAGTACAGCAGCACGAGGGACATGCTGGATGATG ATGGGGACACCACCATGAGCCTGCATTCTCAAGCCTCTGCCACAACTCGGCATCCAGAGCCC CGGCGCACAGAGCACAGGGCTCCCTCTTCAACGTGGCGACCAGTGGCCCTGACCCTGCTGAC TTTGTGCTTGGTGCTGATAGGGCTGGCAGCCCTGGGGCTTTTGTTTTTTCAGTACTACC AGCTCTCCAATACTGGTCAAGACACCATTTCTCAAATGGAAGAAGATTAGGAAATACGTCC CAAGAGTTGCAATCTCTTCAAGTCCAGAATATAAAGCTTGCAGGAAGTCTGCAGCATGTGGC TGAAAAACTCTGTCGTGAGCTGTATAACAAAGCTGGAGCACACAGGTGCAGCCCTTGTACAG AACAATGGAAATGGCATGGAGACAATTGCTACCAGTTCTATAAAGACAGCAAAAGTTGGGAG CCTGGAATTTGCCGCGTCTCAGAGCTACTCTGAGTTTTTCTACTCTTATTGGACAGGGCTTT TGCGCCCTGACAGTGGCAAGGCCTGGCTGTGGATGGAACCCCTTTCACTTCTGAACTG TTCCATATTATAATAGATGTCACCAGCCCAAGAAGCAGAGACTGTGTGGCCATCCTCAATGG GATGATCTTCTCAAAGGACTGCAAAGAATTGAAGCGTTGTGTCTGTGAGAGAAGGGCAGGAA ${\tt TGGTGAAGCCAGAGAGCCTCCATGTCCCCCCTGAAACATTAGGCGAAGGTGAC{\tt TGA}{\tt TTCGCC}}$ CTCTGCAACTACAAATAGCAGAGTGAGCCAGGCGGTGCCAAAGCAAGGGCTAGTTGAGACAT TGGGAAATGGAACATAATCAGGAAAGACTATCTCTCTGACTAGTACAAAATGGGTTCTCGTG TTTCCTGTTCAGGATCACCAGCATTTCTGAGCTTGGGTTTATGCACGTATTTAACAGTCACA AGAAGTCTTATTTACATGCCACCAACCAACCTCAGAAACCCATAATGTCATCTGCCTTCTTG GCTTAGAGATAACTTTTAGCTCTCTTTCTTCTCAATGTCTAATATCACCTCCCTGTTTTCAT GTCTTCCTTACACTTGGTGGAATAAGAAACTTTTTGAAGTAGAGGAAATACATTGAGGTAAC ATCCTTTTCTCTGACAGTCAAGTAGTCCATCAGAAATTGGCAGTCACTTCCCAGATTGTACC AGCAAATACACAAGGAATTCTTTTTGTTTGTTTCAGTTCATACTAGTCCCTTCCCAATCCAT CAGTAAAGACCCCATCTGCCTTGTCCATGCCGTTTCCCAACAGGGATGTCACTTGATATGAG AATCTCAAATCTCAATGCCTTATAAGCATTCCTTCCTGTGTCCATTAAGACTCTGATAATTG TCTCCCCTCCATAGGAATTTCTCCCAGGAAAGAAATATATCCCCATCTCCGTTTCATATCAG AACTACCGTCCCCGATATTCCCTTCAGAGAGATTAAAGACCAGAAAAAAGTGAGCCTCTTCA TCTGCACCTGTAATAGTTTCAGTTCCTATTTTCTTCCATTGACCCATATTTATACCTTTCAG GTACTGAAGATTTAATAATAATAAATGTAAATACTGTGAAAAA

FIGURE 230

MQAKYSSTRDMLDDDGDTTMSLHSQASATTRHPEPRRTEHRAPSSTWRPVALTLLTLCLVLL IGLAALGLLFFQYYQLSNTGQDTISQMEERLGNTSQELQSLQVQNIKLAGSLQHVAEKLCRE LYNKAGAHRCSPCTEQWKWHGDNCYQFYKDSKSWEDCKYFCLSENSTMLKINKQEDLEFAAS QSYSEFFYSYWTGLLRPDSGKAWLWMDGTPFTSELFHIIIDVTSPRSRDCVAILNGMIFSKD CKELKRCVCERRAGMVKPESLHVPPETLGEGD

FIGURE 231

FIGURE 232

GCCGAGCGCAAGAACCCTGCGCAGCCCAGAGCAGCTGCTGGAGGGGAATCGAGGCGCGCTC CCCGAGCCCTCCGGATCCGCCCCCTCCCGGTCCCGCCCCCTCGGAGACTCCTCTGGCTGCT CTGGGGGTTCGCCGGGGCCGGGGACCCGCGGTCCGGGCGCCATCGCTGCTGCTGCTG TCGGTGCTGCGGCCCGCAGGGCCCGTGGCCGTGGGCATCTCCCTGGGCCTTCACCCTGAGCCT AGCTGCCGCCGCGCGCAACACCAACGCGGCGCGCCCGACTCGGTGCAGCCCGGAGCG GAGCGCGAGAAGCCCGGGGCCGAAGGCGCCGGGGAGAATTGGGAGCCGCGCGTCTTGCC CTACCACCCTGCACAGCCCGGCCAGGCCGCCAAAAAGGCCGTCAGGACCCGCTACATCAGCA CGGAGCTGGGCATCAGGCAGAGGCTGCTGGTGGCGGTGCTGACCTCTCAGACCACGCTGCCC ACGCTGGGCGTGGCCGTGAACCGCACGCTGGGGCACCGGCTGGAGCGTGTGGTGTTCCTGAC GGGCGCACGGGCCCGGGCCCCACCTGGCATGGCAGTGACGCTGGGCGAGGAGCGAC CCATTGGACACCTGCACCTGCGCGCTGCGCCACCTGCTGGAGCACGCCGACGACTTTGAC TGGTTCTTCCTGGTGCCTGACACCACCTACACCGAGGCGCACGGCCTGGCACGCCTAACTGG GAGAGCCCACCCCGGCCGCTACTGCCACGGAGGCTTTGGGGGTGCTGCTGTCGCGCATGCTG $\verb|CTGCAACAACTGCGCCCCCACCTGGAAGGCTGCCGCAACGACATCGTCAGTGCGCCCCTGA||$ CGAGTGGCTGGGTCGCTGCATTCTCGATGCCACCGGGGTGGGCTGCACTGGTGACCACGAGG GGGTGCACTATAGCCATCTGGAGCTGAGCCCTGGGGAGCCAGTGCAGGAGGGGGACCCTCAT TTCCGAAGTGCCCTGACAGCCCACCCTGTGCGTGACCCTGTGCACATGTACCAGCTGCACAA AGCTTTCGCCCGAGCTGAACTGGAACGCACGTACCAGGAGATCCAGGAGTTACAGTGGGAGA TCCAGAATACCAGCCATCTGGCCGTTGATGGGGACCGGGCAGCTGCTTGGCCCGTGGGTATT CCAGCACCATCCCGCCCGGCCTCCCGCTTTGAGGTGCTGCGCTGGGACTACTTCACGGAGCA GCACGCTTTCTCCTGCGCCGATGGCTCACCCCGCTGCCCACTGCGTGGGGCTGACCGGGCTG ATGTGGCCGATGTTCTGGGGACAGCTCTAGAGGAGCTGAACCGCCGCTACCACCCGGCCTTG CGGCTCCAGAAGCAGCTGGTGAATGGCTACCGACGCTTTGATCCGGCCCGGGGTATGGA ATACACGCTGGACTTGCAGCTGGAGGCACTGACCCCCCAGGGAGGCCGCCGGCCCCTCACTC GCCGAGTGCAGCTGCTCCGGCCGCTGAGCCGCGTGGAGATCTTGCCTGTGCCCTATGTCACT GAGGCCTCACGTCTCACTGTGCTGCCTCTAGCTGCGGCTGAGCGTGACCTGGCCCCTGG CTTCTTGGAGGCCTTTGCCACTGCAGCACTGGAGCCTGGTGATGCTGCGGCAGCCCTGACCC TGCTGCTACTGTATGAGCCGCGCCAGGCCCAGCGCGTGGCCCATGCAGATGTCTTCGCACCT GTCAAGGCCCACGTGGCAGAGCTGGAGCGGCGTTTCCCCGGTGCCCGGGTGCCATGGCTCAG TGTGCAGACAGCCGCACCCTCACCACTGCGCCTCATGGATCTACTCTCCAAGAAGCACCCGC TGGACACACTGTTCCTGCTGGCCGGGCCAGACACGGTGCTCACGCCTGACTTCCTGAACCGC TGCCGCATGCATGCCATCTCCGGCTGGCAGGCCTTCTTTCCCATGCATTTCCAAGCCTTCCA CCCAGGTGTGGCCCCACCACAAGGGCCTGGGCCCCCAGAGCTGGGCCGTGACACTGGCCGCT TTGATCGCCAGGCAGCCAGCGAGGCCTGCTTCTACAACTCCGACTACGTGGCAGCCCGTGGG CGCCTGGCGGCAGCCTCAGAACAAGAAGAGAGCTGCTGGAGAGCCTGGATGTGTACGAGCT ACCGGGCCCAGACGTGCAGCGCGAGGCTCAGTGAGGACCTGTACCACCGCTGCCTCCAGAGC GTGCTTGAGGGCCTCGGCTCCCGAACCCAGCTGGCCATGCTACTCTTTGAACAGGAGCAGGG CAACAGCACCTGACCCCTGTCCCCGTGGGCCGTGGCATGGCCACACCCCACCCCACTT CTCCCCAAAACCAGAGCCACCTGCCAGCCTCGCTGGGCAGGGCTGGCCGTAGCCAGACCCC AAGCTGGCCCACTGGTCCCCTCTCTGGCTCTGTGGGTCCCTGGGCTCTGGACAAGCACTGGG GGACGTGCCCCAGAGCCACCCACTTCTCATCCCAAACCCAGTTTCCCTGCCCCCTGACGCT GCTGATTCGGGCTGTGGCCTCCACGTATTTATGCAGTACAGTCTGCCTGACGCCAGCCCTGC CTCTGGGCCCTGGGGCTGGGCTGTAGAAGAGTTGTTGGGGAAGGAGGAGCTGAGGAGGGG GCATCTCCCAACTTCTCCCTTTTGGACCCTGCCGAAGCTCCCTGCCTTTAATAAACTGGCCA AGTGTGGAAAAA

FIGURE 233

MRASLLLSVLRPAGPVAVGISLGFTLSLLSVTWVEEPCGPGPPQPGDSELPPRGNTNAARRP
NSVQPGAEREKPGAGEGAGENWEPRVLPYHPAQPGQAAKKAVRTRYISTELGIRQRLLVAVL
TSQTTLPTLGVAVNRTLGHRLERVVFLTGARGRRAPPGMAVVTLGEERPIGHLHLALRHLLE
QHGDDFDWFFLVPDTTYTEAHGLARLTGHLSLASAAHLYLGRPQDFIGGEPTPGRYCHGGFG
VLLSRMLLQQLRPHLEGCRNDIVSARPDEWLGRCILDATGVGCTGDHEGVHYSHLELSPGEP
VQEGDPHFRSALTAHPVRDPVHMYQLHKAFARAELERTYQEIQELQWEIQNTSHLAVDGDRA
AAWPVGIPAPSRPASRFEVLRWDYFTEQHAFSCADGSPRCPLRGADRADVADVLGTALEELN
RRYHPALRLQKQQLVNGYRRFDPARGMEYTLDLQLEALTPQGGRRPLTRRVQLLRPLSRVEI
LPVPYVTEASRLTVLLPLAAAERDLAPGFLEAFATAALEPGDAAAALTLLLLYEPRQAQRVA
HADVFAPVKAHVAELERRFPGARVPWLSVQTAAPSPLRLMDLLSKKHPLDTLFLLAGPDTVL
TPDFLNRCRMHAISGWQAFFPMHFQAFHPGVAPPQGPGPPELGRDTGRFDRQAASEACFYNS
DYVAARGRLAAASEQEEELLESLDVYELFLHFSSLHVLRAVEPALLQRYRAQTCSARLSEDL
YHRCLQSVLEGLGSRTQLAMLLFEQEQGNST

FIGURE 234

GCTCTGGCCGGCCCGGCGATTGGTCACCGCCCGCTAGGGGACAGCCCTGGCCTCCTCTGAT TGGCAAGCGCTGGCCACCCCCACACCCCTTGCGAACGCTCCCCTAGTGGAGAAAAGGAGT AGCTATTAGCCAATTCGGCAGGGCCCGCTTTTTAGAAGCTTGATTTCCTTTGAAGATGAAAG ACTAGCGGAAGCTCTGCCTCTTTCCCCAGTGGGCGAGGGAACTCGGGGCGATTGGCTGGGAA CTGTATCCACCCAAATGTCACCGATTTCTTCCTATGCAGGAAATGAGCAGACCCATCAATAA GAAATTTCTCAGCCTGGCCGAAAATGGTTGGCCCCACGAAGCCACGACAACTGGAGGCAAAG AGGGTTGCTCAACGCCCCGCCTCATTGGAAAACCAAATCAGATCTGGGACCTATATAGCGTG GCGGAGCGGGGCGATGATTGTCGCGCTCGCACCCACTGCAGCTGCGCACAGTCGCATTTCT $\tt TTCCCCGCCCTGAGACCCTGCAGCACCATCTGTC \textbf{ATG} \tt CGGCTGGGCTGTTTGGTTTGAGC$ GCTCGCCGTCTTTTGGCGGCAGCGGCGACGCGAGGGCTCCCGGCCGCCCCGCGTCCGCTGGGA ATCTAGCTTCTCCAGGACTGTGGTCGCCCCGTCCGCTGTGGCGGGAAAGCGGCCCCCAGAAC CGACCACACCGTGGCAAGAGGACCCAGAACCCGAGGACGAAAACTTGTATGAGAAGAACCCA GACTCCCATGGTTATGACAAGGACCCCGTTTTGGACGTCTGGAACATGCGACTTGTCTTCTT CTTTGGCGTCTCCATCATCCTGGTCCTTGGCAGCACCTTTGTGGCCTATCTGCCTGACTACA GGATGAAAGAGTGGTCCCGCCGCGAAGCTGAGAGGCTTGTGAAATACCGAGAGGCCAATGGC CTTCCCATCATGGAATCCAACTGCTTCGACCCCAGCAAGATCCAGCTGCCAGAGGATGAGTG CTCTTCTCAGAGCACCTAATTAAAGGGGCTGAAAGTCTGAA

FIGURE 235

MAAGLFGLSARRLLAAAATRGLPAARVRWESSFSRTVVAPSAVAGKRPPEPTTPWQEDPEPE DENLYEKNPDSHGYDKDPVLDVWNMRLVFFFGVSIILVLGSTFVAYLPDYRMKEWSRREAER LVKYREANGLPIMESNCFDPSKIQLPEDE

FIGURE 236

FIGURE 237

GCGCGGCTATGCCGCTTGCTCGTCCTGTTGCTCCTGGGGCCCGGCGGCTGGTGCCT TGCAGAACCCCCACGCGACAGCCTGCGGGAGGAACTTGTCATCACCCCGCTGCCTTCCGGGG ACGTAGCCGCCACATTCCAGTTCCGCACGCGCTGGGATTCGGAGCTTCAGCGGGAAGGAGTG TCCCATTACAGGCTCTTTCCCAAAGCCCTGGGGCAGCTGATCTCCAAGTATTCTCTACGGGA GCTGCACCTGTCATTCACACAAGGCTTTTGGAGGACCCGATACTGGGGGCCCACCCTTCCTGC AGGCCCCATCAGGTGCAGAGCTGTGGGTCTGGTTCCAAGACACTGTCACTGATGTGGATAAA TCTTGGAAGGAGCTCAGTAATGTCCTCTCAGGGATCTTCTGCGCCTCTCTCAACTTCATCGA CTCCACCAACACAGTCACTCCCACTGCCTCCTTCAAACCCCTGGGTCTGGCCAATGACACTG ACCACTACTTTCTGCGCTATGCTGTGCTGCCGCGGGAGGTGGTCTGCACCGAAAACCTCACC CCCTGGAAGAGCTCTTGCCCTGTAGTTCCAAGGCAGGCCTCTCTGTGCTGCTGAAGGCAGA TCGCTTGTTCCACACCAGCTACCACTCCCAGGCAGTGCATATCCGCCCTGTTTGCAGAAATG CACGCTGTACTAGCATCTCCTGGGAGCTGAGGCAGACCCTGTCAGTTGTATTTGATGCCTTC ATCACGGGGCAGGGAAAGAAGACTGGTCCCTCTTCCGGATGTTCTCCCGAACCCTCACGGA GCCCTGCCCCTGGCTTCAGAGAGCCGAGTCTATGTGGACATCACCACCTACAACCAGGACA ACGAGACATTAGAGGTGCACCCACCCCGACCACTACATATCAGGACGTCATCCTAGGCACT CGGAAGACCTATGCCATCTATGACTTGCTTGACACCGCCATGATCAACAACTCTCGAAACCT CAACATCCAGCTCAAGTGGAAGAGACCCCCAGAGAATGAGGCCCCCCAGTGCCCTTCCTGC ATGCCCAGCGGTACGTGAGTGGCTATGGGCTGCAGAAGGGGGAGCTGAGCACACTGCTGTAC AACACCCACCCATACCGGGCCTTCCCGGTGCTGCTGGACACCGTACCCTGGTATCTGCG GCTGTATGTGCACACCCTCACCATCACCTCCAAGGGCAAGGAACAAACCAAGTTACATCC ACTACCAGCCTGCCCAGGACCGGCTGCAACCCCACCTCCTGGAGATGCTGATTCAGCTGCCG GCCAACTCAGTCACCAAGGTTTCCATCCAGTTTGAGCGGGCGCTGCTGAAGTGGACCGAGTA CACGCCAGATCCTAACCATGGCTTCTATGTCAGCCCATCTGTCCTCAGCGCCCTTGTGCCCA GCATGGTAGCAGCCAAGCCAGTGGACTGGGAAGAGAGTCCCCTCTTCAACAGCCTGTTCCCA GTCTCTGATGGCTCTAACTACTTTGTGCGGCTCTACACGGAGCCGCTGCTGGTGAACCTGCC GACACCGGACTTCAGCATGCCCTACAACGTGATCTGCCTCACGTGCACTGTGGTGGCCGTGT GCTACGGCTCCTTCTACAATCTCCTCACCCGAACCTTCCACATCGAGGAGCCCCGCACAGGT GGCCTGGCCAAGCGGCTGGCCAACCTTATCCGGCGCGCCCGAGGTGTCCCCCCACTC**TGA**TT CTTGCCCTTTCCAGCAGCTGCAGCTGCCGTTTCTCTCTGGGGAGGGGAGCCCAAGGGCTGTT TCTGCCACTTGCTCTCAGAGTTGGCTTTTGAACCAAAGTGCCCTGGACCAGGTCAGGGC CTACAGCTGTGTTGTCCAGTACAGGAGCCACGAGCCAAATGTGGCATTTGAATTTGAATTAA CTTAGAAATTCATTTCCTCACCTGTAGTGGCCACCTCTATATTGAGGTGCTCAATAAGCAAA AGTGGTCGGTGGCTGTATTGGACAGCACAGAAAAAGATTTCCATCACCACAGAAAGGTC GGCTGGCAGCACTGGCCAAGGTGATGGGGTGTGCTACACAGTGTATGTCACTGTAGTGGA

FIGURE 238

MPLALLVLLLLGPGGWCLAEPPRDSLREELVITPLPSGDVAATFQFRTRWDSELQREGVSHY
RLFPKALGQLISKYSLRELHLSFTQGFWRTRYWGPPFLQAPSGAELWVWFQDTVTDVDKSWK
ELSNVLSGIFCASLNFIDSTNTVTPTASFKPLGLANDTDHYFLRYAVLPREVVCTENLTPWK
KLLPCSSKAGLSVLLKADRLFHTSYHSQAVHIRPVCRNARCTSISWELRQTLSVVFDAFITG
QGKKDWSLFRMFSRTLTEPCPLASESRVYVDITTYNQDNETLEVHPPPTTTYQDVILGTRKT
YAIYDLLDTAMINNSRNLNIQLKWKRPPENEAPPVPFLHAQRYVSGYGLQKGELSTLLYNTH
PYRAFPVLLLDTVPWYLRLYVHTLTITSKGKENKPSYIHYQPAQDRLQPHLLEMLIQLPANS
VTKVSIQFERALLKWTEYTPDPNHGFYVSPSVLSALVPSMVAAKPVDWEESPLFNSLFPVSD
GSNYFVRLYTEPLLVNLPTPDFSMPYNVICLTCTVVAVCYGSFYNLLTRTFHIEEPRTGGLA
KRLANLIRRARGVPPL

FIGURE 239

FIGURE 240

MGSSSFLVLMVSLVLVTLVAVEGVKEGIEKAGVCPADNVRCFKSDPPQCHTDQDCLGERKCC YLHCGFKCVIPVKELEEGGNKDEDVSRPYPEPGWEAKCPGSSSTRCPQK

Signal sequence:

amino acids 1-19

N-myristoylation sites:

amino acids 23-29, 27-33, 32-38, 102-108

WAP-type 'four-disulfide core' domain signature:

amino acids 49-63

FIGURE 241

AAACTCAGCACTTGCCGGAGTGGCTCATTGTTAAGACAAAGGGTGTGCACTTCCTGGCCAGG AAACCTGAGCGGTGAGACTCCCAGCTGCCTACATCAAGGCCCCAGGACATGCAGAACCTTCC TCTAGAACCCGACCACCACCATGAGGTCCTGCCTGTGGAGATGCAGGCACCTGAGCCAAGG CGTCCAGTGGTCCTTGCTTCTGGCTGTCCTGGTCTTCTTCTCTTCGCCTTGCCCTCTTTTA CTACAGTCCCTGGCAAAGCCTAAGTCCCAGGCACCACAAGGGCGAGGAGGACAACCATCTA TGCAGAGCCAGCGCCAGAGAACAATGCCCTCAACACACAAACCCAGCCCAAGGCCCACACCA ACAGCACAGAGGGCAGCATGGAAGAGCCCAGAAAAAGAGAAAACCATGGTGAACACTGTC ACCCAGAGGGCAAGATGCAGGGATGGCCTCTGGCAGGACAGAGGCACAATCATGGAAGAGCC AGGACACAAAGACGACCCAAGGAAATGGGGGCCAGACCAGGAAGCTGACGGCCTCCAGGACG GTGTCAGAGAAGCACCAGGGCAAAGCGGCAACCACAGCCAAGACGCTCATTCCCAAAAGTCA GCACAGAATGCTGGCTCCCACAGGAGCAGTGTCAACAAGGACGAGACAGAAAGGAGTGACCA CAGCAGTCATCCCACCTAAGGAGAAGAAACCTCAGGCCACCCCACCCCCTGCCCCTTTCCAG AGCCCCACGACGCAGAGAAACCAAAGACTGAAGGCCGCCAACTTCAAATCTGAGCCTCGGTG GGATTTTGAGGAAAAATACAGCTTCGAAATAGGAGGCCTTCAGACGACTTGCCCTGACTCTG TGAAGATCAAAGCCTCCAAGTCGCTGTGGCTCCAGAAACTCTTTCTGCCCAACCTCACTCTC TTCCTGGACTCCAGACACTTCAACCAGAGTGAGTGGGACCGCCTGGAACACTTTGCACCACC CTTTGGCTTCATGGAGCTCAACTACTCCTTGGTGCAGAAGGTCGTGACACGCTTCCCTCCAG TGCCCCAGCAGCTGCTCCTGGCCAGCCTCCCCGCTGGGAGCCTCCGGTGCATCACCTGT GCCGTGGTGGCCAACGGGGCCATCCTGAACAACTCCCACATGGGCCAGGAGATAGACAGTCA CGACTACGTGTTCCGATTGAGCGGAGCTCTCATTAAAGGCTACGAACAGGATGTGGGGACTC GGACATCCTTCTACGGCTTTACCGCCTTCTCCCTGACCCAGTCACTCCTTATATTGGGCAAT CGGGGTTTCAAGAACGTGCCTCTTGGGAAGGACGTCCGCTACTTGCACTTCCTGGAAGGCAC CCGGGACTATGAGTGGCTGGAAGCACTGCTTATGAATCAGACGGTGATGTCAAAAAACCTTT TCTGGTTCAGGCACAGACCCCAGGAAGCTTTTCGGGAAGCCCTGCACATGGACAGGTACCTG TTGCTGCACCCAGACTTTCTCCGATACATGAAGAACAGGTTTCTGAGGTCTAAGACCCTGGA TGGTGCCCACTGGAGGATATACCGCCCCACCACTGGGGCCCTCCTGCTGCTCACTGCCCTTC AGCTCTGTGACCAGGTGAGTGCTTATGGCTTCATCACTGAGGGCCCATGAGCGCTTTTCTGAT CACTACTATGATACATCATGGAAGCGGCTGATCTTTTACATAAACCATGACTTCAAGCTGGA GAGAGAAGTCTGGAAGCGGCTACACGATGAAGGGATAATCCGGCTGTACCAGCGTCCTGGTC CCGGAACTGCCAAAGCCAAGAAC**TGA**CCGGGGCCAGGGCTGCCATGGTCTCCTTGCCTGCTC CAAGGCACAGGATACAGTGGGAATCTTGAGACTCTTTGGCCATTTCCCATGGCTCAGACTAA GCTCCAAGCCCTTCAGGAGTTCCAAGGGAACACTTGAACCATGGACAAGACTCTCTCAAGAT GGCAAATGGCTAATTGAGGTTCTGAAGTTCTTCAGTACATTGCTGTAGGTCCTGAGGCCAGG GATTTTTAATTAAATGGGGTGATGGGTGGCCAATACCACAATTCCTGCTGAAAAACACTCTT CCAGTCCAAAAGCTTCTTGATACAGAAAAAAGAGCCTGGATTTACAGAAACATATAGATCTG GTTTGAATTCCAGATCGAGTTTACAGTTGTGAAATCTTGAAGGTATTACTTAACTTCACTAC AGATTGTCTAGAAGACCTTTCTAGGAGTTATCTGATTCTAGAAGGGTCTATACTTGTCCTTG TCTTTAAGCTATTTGACAACTCTACGTGTTGTAGAAAACTGATAATAATACAAATGATTGTT

FIGURE 242

MRSCLWRCRHLSQGVQWSLLLAVLVFFLFALPSFIKEPQTKPSRHQRTENIKERSLQSLAKP
KSQAPTRARRTTIYAEPAPENNALNTQTQPKAHTTGDRGKEANQAPPEEQDKVPHTAQRAAW
KSPEKEKTMVNTLSPRGQDAGMASGRTEAQSWKSQDTKTTQGNGGQTRKLTASRTVSEKHQG
KAATTAKTLIPKSQHRMLAPTGAVSTRTRQKGVTTAVIPPKEKKPQATPPPAPFQSPTTQRN
QRLKAANFKSEPRWDFEEKYSFEIGGLQTTCPDSVKIKASKSLWLQKLFLPNLTLFLDSRHF
NQSEWDRLEHFAPPFGFMELNYSLVQKVVTRFPPVPQQQLLLASLPAGSLRCITCAVVGNGG
ILNNSHMGQEIDSHDYVFRLSGALIKGYEQDVGTRTSFYGFTAFSLTQSLLILGNRGFKNVP
LGKDVRYLHFLEGTRDYEWLEALLMNQTVMSKNLFWFRHRPQEAFREALHMDRYLLLHPDFL
RYMKNRFLRSKTLDGAHWRIYRPTTGALLLLTALQLCDQVSAYGFITEGHERFSDHYYDTSW
KRLIFYINHDFKLEREVWKRLHDEGIIRLYQRPGPGTAKAKN

Cytoplasmic Domain:

amino acids 1-10

Type II Transmembrane Domain:

amino acids 11-35

Lumenal catalytic Domain:

amino acids 36-600

Ribonucleotide Reductase small subunit Signature:

amino acids 481-496

N-glycosylation Sites:

amino acids 300-303, 311-314, 331-334, 375-378, 460-463

FIGURE 243

FIGURE 244

MRGPGHPLLLGLLLVLGPSPEQRVEIVPRDLRMKDKFLKHLTGPLYFSPKCSKHFHRLYHNT RDCTIPAYYKRCARLLTRLAVSPVCMEDK

FIGURE 245

GGGCTGGGCCCCGCCGCAGCTCCAGCTGGCCGGCTTGGTCCTGCGGTCCCTTCTCTGGGAGG ${\tt CCCGACCCGGCCGCGCCCAGCCCCACC}$ CTCACCGCAATCGCTCTGTTGGTGCTGGGGGGCTCCCCTGGTGCTGGCCGGCGAGGACTGCCT GTGGTACCTGGACCGGAATGGCTCCTGGCATCCGGGGTTTAACTGCGAGTTCTTCACCTTCT GCTGCGGGACCTGCTACCATCGGTACTGCTGCAGGGACCTGACCTTGCTTATCACCGAGAGG CAGCAGAAGCACTGCCTGGCCTTCAGCCCCAAGACCATAGCAGGCATCGCCTCAGCTGTGAT $\verb|CCTCTTTGTTGCTGTTGCTACCACCACCATCTGCTGCTTCCTGTTCCTGTTGCTACCTGT|\\$ ACCGCCGGCGCCAGCAGCTCCAGAGCCCATTTGAAGGCCAGGAGATTCCAATGACAGGCATC CCAGTGCAGCCAGTATACCCATACCCCCAGGACCCCAAAGCTGGCCCTGCACCCCCACAGCC TGGCTTCATGTACCCACCTAGTGGTCCTGCTCCCCAATATCCACTCTACCCAGCTGGGCCCC CAGTCTACAACCCTGCAGCTCCTCCTCCTATATGCCACCACAGCCCTCTTACCCGGGAGCC TGAGGAACCAGCCATGTCTCTGCTGCCCCTTCAGTGATGCCAACCTTGGGAGATGCCCTCAT CCTGTACCTGCATCTGGTCCTGGGGGTGGCAGGAGTCCTCCAGCCACCAGGCCCCAGACCAA GCCAAGCCCTGGGCCCTACTGGGGACAGAGCCCCAGGGAAGTGGAACAGGAGCTGAACTAGA ACTATGAGGGGTTGGGGGGGGGCTTGGAATTATGGGCTATTTTTACTGGGGGCAAGGGAGG GAGATGACAGCCTGGGTCACAGTGCCTGTTTTCAAATAGTCCCTCTGCTCCCAAGATCCCAG CCAGGAAGGCTGGGGCCCTACTGTTTGTCCCCTCTGGGCTGGGGTGGGGGGAGGAGGAGGT TCCGTCAGCAGCTGGCCAGTAGCCCTCTCTCTGGCTGCCCCACTGGCCACATCTCTGGCCTG CTAGATTAAAGCTGTAAAGACAAAA

FIGURE 246

MPPAGLRRAAPLTAIALLVLGAPLVLAGEDCLWYLDRNGSWHPGFNCEFFTFCCGTCYHRYC CRDLTLLITERQQKHCLAFSPKTIAGIASAVILFVAVVATTICCFLCSCCYLYRRRQQLQSP FEGQEIPMTGIPVQPVYPYPQDPKAGPAPPQPGFMYPPSGPAPQYPLYPAGPPVYNPAAPPP YMPPQPSYPGA

Transmembrane Domains:

amino acids 10-28, 85-110

N-glycosylation Site:

amino acids 38-41

N-myristoylation Sites:

amino acids 5-10, 88-93

FIGURE 247

GGGGGAGCTAGGCCGGCGGCAGTGGTGGTGGCGGCGCGCAAGGGTGAGGGCGGCCCCAGAA CCCCAGGTAGGTAGAGCAAGAAG**ATG**GTGTTTCTGCCCCTCAAATGGTCCCTTGCAACCATG TCATTTCTACTTTCCTCACTGTTGGCTCTCTTAACTGTGTCCACTCCTTCATGGTGTCAGAG TCATTTCTACTTTCCTCACTGTTGGCTCTTTAACTGTGTCCACTCCTTCATGGTGTCAGAG
CACTGAAGCATCTCCAAAACGTAGTGATGGGACACCATTTCCTTGGAATAAAATACGACTTC
CTGAGTACGTCATCCCAGTTCATTATGATCTCTTGATCCATGCAAACCTTACCACGCTGACC
TTCTGGGGAACCACGAAAGTAGAAATCACAGCCAGTCAGCCCACCAGCACCATCATCCTGCA
TAGTCACCACCTGCAGATATCTAGGGCCACCCTCAGGAAGGGAGCTGGAGAGAGCCTATCGG
AAGAACCCCTGCAGGTCCTGGAACACCCCCCCTCAGGAGCAAATTGCACTGCTGCCCGAG CCCCTCCTTGTCGGGCTCCCGTACACAGTTGTCATTCACTATGCTGGCAATCTTTCGGAGAC AGATGAGCACCTATCTGGTGGCCTTCATCATTTCAGATTTTGAGTCTGTCAGCAAGATAACC AGATGAGCACCTATCTGGTGGCCTTCATCATTTCAGATTTTGAGTCTGTCAGCAAGATAACC
AAGAGTGGAGTCAAGGTTTCTGTTTATGCTGTGCCAGACAAGATAAATCAAGCAGATTATGC
ACTGGATGCTGCGGTGACTCTTCTAGAATTTTATGAGGATTATTTCAGCATACCGTATCCCC
TACCCAAACAAGATCTTGCTGCTATTCCCGACTTTCAGTCTGGTGCTATTGGAAAACTGGGGA
CTGACAACATATAGAGAATCTGCTCTGTTTGTTTGATGCAGAAAAGTCTTCTGCATCAAGTAA
GCTTGGCATCACAGTGACTGGCCCCATGAACTGGCCCACCAGTGGTTTTGGGAACCTGGTCA CTATGGAATGGTGGAATGATCTTTGGCTAAATGAAGGATTTGCCAAATTTATGGAGTTTGTG TCTGTCAGTGTGACCCATCCTGAACTGAAAGTTGGAGATTATTTCTTTGGCAAATGTTTTGA CGCAATGGAGGTAGATGCTTTAAATTCCTCACACCCTGTGTCTACACCTGTGGAAAATCCTG CTCAGATCCGGGAGATGTTTGATGATGTTTCTTATGATAAGGGAGCTTGTATTCTGAATATG CTAAGGGAGTATCTTAGCGCTGACGCATTTAAAAGTGGTATTGTACAGTATCTCCAGAAGCA TAGCTATAAAAATACAAAAAACGAGGACCTGTGGGATAGTATGGCAAGTATTTGCCCTACAG ATGGTGTAAAAGGGATGGATGGCTTTTGCTCTAGAAGTCAACATTCATCTTCATCCTCACAT TGGCATCAGGAAGGGGTGGATGTGAAAACCATGATGAACACTTGGACACTGCAGAGGGGTTT TCCCCTAATAACCATCACAGTGAGGGGGGGGGAATGTACACATGAAGCAAGAGCACTACATGA AGGGCTCTGACGGCCCCCGGACACTGGGTACCTGTGGCATGTTCCATTGACATTCATCACC AGCAAATCCAACATGGTCCATCGATTTTTGCTAAAAACAAAAACAGATGTGCTCATCCTCCC TTCTCAGCTCCGTTGTGTCCAACAGACAATTGAAACCATTGAAGAAAACATCGGTTGGATGG ATAAGAATTTTGATAAAATCAGAGTGTGGCTGCAAAGTGAAAAGCTTGAACGTATG**TAA**AAA TTCCTCCCTTGCCCGGTTCCTGTTATCTCTAATCACCAACATTTTGTTGAGTGTATTTTCAA ACTAGAGATGGCTGTTTTGGCTCCAACTGGAGATACTTTTTTCCCTTCAACTCATTTTTTGA CTATCCCTGTGAAAAGAATAGCTGTTAGTTTTTCATGAATGGGCTTTTTCATGAATGGGCTA TCGCTACCATGTGTTTTGTTCATCACAGGTGTTGCCCTGCAACGTAAACCCAAGTGTTGGGT

FIGURE 248

MVFLPLKWSLATMSFLLSSLLALLTVSTPSWCQSTEASPKRSDGTPFPWNKIRLPEYVIPVH YDLLIHANLTTLTFWGTTKVEITASQPTSTIILHSHHLQISRATLRKGAGERLSEEPLQVLE HPPQEQIALLAPEPLLVGLPYTVVIHYAGNLSETFHGFYKSTYRTKEGELRILASTQFEPTA ARMAFPCFDEPAFKASFSIKIRREPRHLAISNMPLVKSVTVAEGLIEDHFDVTVKMSTYLVA FIISDFESVSKITKSGVKVSVYAVPDKINQADYALDAAVTLLEFYEDYFSIPYPLPKQDLAA IPDFQSGAMENWGLTTYRESALLFDAEKSSASSKLGITVTVAHELAHQWFGNLVTMEWWNDL WLNEGFAKFMEFVSVSVTHPELKVGDYFFGKCFDAMEVDALNSSHPVSTPVENPAQIREMFD DVSYDKGACILNMLREYLSADAFKSGIVQYLQKHSYKNTKNEDLWDSMASICPTDGVKGMDG FCSRSOHSSSSSHWHOEGVDVKTMMNTWTLQRGFPLITITVRGRNVHMKQEHYMKGSDGAPD TGYLWHVPLTFITSKSNMVHRFLLKTKTDVLILPEEVEWIKFNVGMNGYYIVHYEDDGWDSL TGLLKGTHTAVSSNDRASLINNAFQLVSIGKLSIEKALDLSLYLKHETEIMPVFQGLNELIP MYKLMEKRDMNEVETQFKAFLIRLLRDLIDKQTWTDEGSVSEQMLRSELLLLACVHNYQPCV QRAEGYFRKWKESNGNLSLPVDVTLAVFAVGAQSTEGWDFLYSKYQFSLSSTEKSQIEFALC RTONKEKLOWLLDESFKGDKIKTOEFPOILTLIGRNPVGYPLAWQFLRKNWNKLVQKFELGS SSIAHMVMGTTNQFSTRTRLEEVKGFFSSLKENGSQLRCVQQTIETIEENIGWMDKNFDKIR **VWLQSEKLERM**

Signal peptide:

amino acids 1-34

N-glycosylation sites:

amino acids 70-74, 154-158, 414-418, 760-764, 901-905

Neutral zinc metallopeptidases, zinc-binding region signature:

amino acids 350-360

FIGURE 249

CAGCCACAGACGGGTC**ATG**AGCGCGGTATTACTGCTGGCCCTCCTGGGGTTCATCCTCCCAC TGCCAGGAGTGCAGGCGCTGCTCTGCCAGTTTGGGACAGTTCAGCATGTGTGGAAGGTGTCC GACCTACCCCGGCAATGGACCCCTAAGAACACCAGCTGCGACAGCGGCTTGGGGTGCCAGGA CACGTTGATGCTCATTGAGAGCGGACCCCAAGTGAGCCTGGTGCTCTCCAAGGGCTGCACGG AGGCCAAGGACCAGGAGCCCCGCTCACTGAGCACCGGATGGGCCCCGGCCTCTCCCTGATC TTGGGCCCCACAGCCCCCAGCAGACCCAGGATCCTTGAGGTGCCCAGTCTGCTTGTCTATGG AAGGCTGTCTGGAGGGGACAACAGAAGAGATCTGCCCCAAGGGGACCACACACTGTTATGAT CCAGCCAGGTTGCAACCTGCTCAATGGGACACAGGAAATTGGGCCCGTGGGTATGACTGAGA ACTGCAATAGGAAAGATTTTCTGACCTGTCATCGGGGGACCACCATTATGACACACGGAAAC TTGGCTCAAGAACCCACTGATTGGACCACATCGAATACCGAGATGTGCGAGGTGGGGCAGGT GTGTCAGGAGACGCTGCTCATAGATGTAGGACTCACATCAACCCTGGTGGGGACAAAAG GCTGCAGCACTGTTGGGGCTCAAAATTCCCAGAAGACCACCATCCACTCAGCCCCTCCTGGG GTGCTTGTGGCCTCCTATACCCACTTCTGCTCCTCGGACCTGTGCAATAGTGCCAGCAGCAG CAGCGTTCTGCTGAACTCCCTCCTCCTCAAGCTGCCCCTGTCCCAGGAGACCGGCAGTGTC CTACCTGTGTGCAGCCCCTTGGAACCTGTTCAAGTGGCTCCCCCGAATGACCTGCCCCAGG GGCGCCACTCATTGTTATGATGGGTACATTCATCTCTCAGGAGGTGGGCTGTCCACCAAAAT GAGCATTCAGGGCTGCGTGGCCCAACCTTCCAGCTTCTTGTTGAACCACACCAGACAAATCG GGGGCTGAGGCCTGGAGTCTCTCACTTGGGGGGTGGGGCTGGCACTGGCCCCAGCGCTGTG GTGGGGAGTGGTTTGCCCTTCCTGC**TAA**CTCTATTACCCCCACGATTCTTCACCGCTGCTGA CCACCCACACTCAACCTCCTCTGACCTCATAACCTAATGGCCTTGGACACCAGATTCTTTC ACACTGGGGAGAGCCTGGAGCATCCGGACTTGCCCTATGGGAGAGGGGACGCTGGAGGAGTG GCTGCATGTATCTGATAATACAGACCCTGTCCTTTCA

FIGURE 250

MSAVLLLALLGFILPLPGVQALLCQFGTVQHVWKVSDLPRQWTPKNTSCDSGLGCQDTLMLI
ESGPQVSLVLSKGCTEAKDQEPRVTEHRMGPGLSLISYTFVCRQEDFCNNLVNSLPLWAPQP
PADPGSLRCPVCLSMEGCLEGTTEEICPKGTTHCYDGLLRLRGGGIFSNLRVQGCMPQPGCN
LLNGTQEIGPVGMTENCNRKDFLTCHRGTTIMTHGNLAQEPTDWTTSNTEMCEVGQVCQETL
LLIDVGLTSTLVGTKGCSTVGAQNSQKTTIHSAPPGVLVASYTHFCSSDLCNSASSSSVLLN
SLPPQAAPVPGDRQCPTCVQPLGTCSSGSPRMTCPRGATHCYDGYIHLSGGGLSTKMSIQGC
VAQPSSFLLNHTRQIGIFSAREKRDVQPPASQHEGGGAEGLESLTWGVGLALAPALWWGVVC
PSC

FIGURE 251

CAGGATGAGGGGGAATCTGGCCCTGGTGGGCGTTCTAATCAGCCTGGCCTTCCTGTCACTGCTG CCATCTGGACATCCTCAGCCGGCTGGCGATGACGCCTGCTCTGTGCAGATCCTCGTCCCTGG CCTCAAAGGGGATGCGGGAGAGAGAGGGACAAAGGCGCCCCCGGACGGCCTGGAAGAGTCG GCCCCACGGGAGAAAAGGAGACATGGGGGACAAAGGACAGAAAGGCAGTGTGGGTCGTCAT GGAAAAATTGGTCCCATTGGCTCTAAAGGTGAGAAAGGAGATTCCGGTGACATAGGACCCCC AGATGGACAACCAGGTCTCTCAGCTGACCAGCGAGCTCAAGTTCATCAAGAATGCTGTCGCC GGTGTGCGCGAGACGGAGAGCAAGATCTACCTGCTGGTGAAGGAGGAGAAGCGCTACGCGGA CGCCCAGCTGTCCTGCCAGGGCCGCGGGGGCACGCTGAGCATGCCCAAGGACGAGGCTGCCA ATGGCCTGATGGCCGCATACCTGGCGCAAGCCGGCCTGGCCCGTGTCTTCATCGGCATCAAC GACCTGGAGAAGGAGGCGCCTTCGTGTACTCTGACCACTCCCCCATGCGGACCTTCAACAA GTGGCGCAGCGGTGAGCCCAACAATGCCTACGACGAGGAGGACTGCGTGGAGATGGTGGCCT CGGGCGGCTGGAACGACGTGGCCTGCCACACCACCATGTACTTCATGTGTGAGTTTGACAAG ${ t GAGAACATG}{ t TGA}{ t GCCTCAGGCTGGGGCTGCCCATTGGGGGCCCCACATGTCCCTGCAGGGTT}$ GGCAGGGACAGAGCCCAGACCATGGTGCCAGCCAGGGAGCTGTCCCTCTGTGAAGGGTGGAG GCTCACTGAGTAGAGGGCTGTTGTCTAAACTGAGAAAATGGCCTATGCTTAAGAGGAAAAATG AAAGTGTTCCTGGGGTGCTGTCTCTGAAGAAGCAGAGTTTCATTACCTGTATTGTAGCCCCA ATGTCATTATGTAATTATTACCCAGAATTGCTCTTCCATAAAGCTTGTGCCTTTGTCCAAGC

FIGURE 252

MRGNLALVGVLISLAFLSLLPSGHPQPAGDDACSVQILVPGLKGDAGEKGDKGAPGRPGRVG
PTGEKGDMGDKGQKGSVGRHGKIGPIGSKGEKGDSGDIGPPGPNGEPGLPCECSQLRKAIGE
MDNQVSQLTSELKFIKNAVAGVRETESKIYLLVKEEKRYADAQLSCQGRGGTLSMPKDEAAN
GLMAAYLAQAGLARVFIGINDLEKEGAFVYSDHSPMRTFNKWRSGEPNNAYDEEDCVEMVAS
GGWNDVACHTTMYFMCEFDKENM

FIGURE 253

AGTGACTGCAGCCTTCCTAGATCCCCTCCACTCGGTTTCTCTCTTTTGCAGGAGCACCGGCAG CACCAGTGTGTGAGGGGAGCAGCCAGCCTTCCTTGATCCTGCCAGACCACC CAGCCCCGGCACAGAGCTGCTCCACAGGCACC<u>ATG</u>AGGATCATGCTGCTATTCACAGCCAT TTCCTGGCGGGGGCCGCAGCAAGAGGGATCCAGATCTCTACCAGCTGCTCCAGAGACTCTTC AAAAGCCACTCATCTCTGGAGGGATTGCTCAAAGCCCTGAGCCAGGCTAGCACAGATCCTAA GGAATCAACATCTCCCGAGAAACGTGACATGCATGACTTCTTTGTGGGACTTATGGGCAAGA GGAGCGTCCAGCCAGAGGGAAAGACAGGACCTTTCTTACCTTCAGTGAGGGTTCCTCGGCCC CTTCATCCCAATCAGCTTGGATCCACAGGAAAGTCTTCCCTGGGAACAGAGGAGCAGAGACC $\verb|TTTATAA| GACTCTCCTACGGATGTGAATCAAGAGAACGTCCCCAGCTTTGGCATCCTCAAGT|$ ATCCCCCGAGAGCAGAATAGGTACTCCACTTCCGGACTCCTGGACTGCATTAGGAAGACCTC AACATTCTTGTGCTTTGACTCCTTCTCCATCTTTTCTACCTGACCCTGGTGTGGAAACTGCA TAGTGAATATCCCCAACCCCAATGGGCATTGACTGTAGAATACCCTAGAGTTCCTGTAGTGT CCTACATTAAAAATATAATGTCTCTCTCTATTCCTCAACAATAAAGGATTTTTGCATATGAA

FIGURE 254

MRIMLLFTAILAFSLAQSFGAVCKEPQEEVVPGGGRSKRDPDLYQLLQRLFKSHSSLEGLLK ALSQASTDPKESTSPEKRDMHDFFVGLMGKRSVQPEGKTGPFLPSVRVPRPLHPNQLGSTGK SSLGTEEQRPL

Important features:

Signal peptide:

amino acids 1-18

Tyrosine kinase phosphorylation site.

amino acids 36-45

N-myristoylation site.

amino acids 33-39, 59-65

Amidation site.

amino acids 90-94

Leucine zipper pattern.

amino acids 43-65

Tachykinin family signature.

amino acids 86-92

FIGURE 255

GGGCGTCTCCGGCTGCTCCTATTGAGCTGTCTGCTCGCTGTGCCCGCTGTGCCTGTGCC CGCGCTGTCGCCGCTGCTACCGCGTCTGCTGGACGCGGGAGACGCCAGCGAGCTGGTGATTG GAGCCCTGCGGAGAGCTCAAGCGCCCAGCTCTGCCCCAGGAGCCCAGGCTGCCCCGTGAGTC ${\tt CCATAGTTGCTGCAGGAGTGGAGCC}$ AGCTGCGTCCTGGGTGTGTCATCCCCTTGGGGC TGCTGTTCCTGGTCTGCGGATCCCAAGGCTACCTCCTGCCCAACGTCACTCTCTTAGAGGAG CTGCTCAGCAAATACCAGCACAACGAGTCTCACTCCCGGGTCCGCAGAGCCATCCCCAGGGA GGACAAGGAGGAGATCCTCATGCTGCACAACAAGCTTCGGGGCCCAGGTGCAGCCTCAGGCCT CCAACATGGAGTACATGGTGAGCGCCGGCTCCGGCCGCAGAGGCTGGCACCGGGGGTGGGGC TTGAGACAGGGTCTCACTCTGCCACTGACGCTGGAGTGCAATGGCACAATCGTCATGCCCTG AAACCT**TAG**ACTCCCGGGGTTAAGCGATCCTGCTTCAGCCTCCCAAGTAGCTGGAACTACAG GCATGCACCATGGTGCCCAGCTAGATTTTAAATATTTTGTGGAGATGGGGGTCTTGCTACGT TGCCCAGGCTGGTCTTGAACTCCTAGGCTCAAGCAATCCTCCTGCCTCAGCCTCTCAAAGTG CTAGGATTATAGGCATGAGTCACCCTGTCTGGCTCTGGCTCTGTTCTTAACATTCTGCCAAA ACAACACGTGGGTTCCCTGTGCAGAGCCTGCCTCGTTGCCTTCATGTCACTCTTGGTAGC TCCACTGGGAACACAGCTCTCAGCCTTTCCCACCTGGAGGCAGAGTGGGGAGGGGCCCAGGG CTGGGCTTTGCTGATGCTGATCTCAGCTGTGCCACACGCTAGCTGCACCACCCTGACTTCTC GTGAGATAAGTCGAGGCTGTGAAGGGCCCGGCACAGACTGACCTGCCTCCCCAACCCCTAGG CTTTGCTAACCGGGAAAGGAGCTAACGGTGACAGAAGACAGCCAAGGTCAACCCTCCCGGGT GATTGTGATGGGTGTTCCAGGTGTGGTTGGGCGATGCTGCTACTTGACCCCAAGCTCCAGTG TGGAAACTTCCTTCCTGGCTGGTTTTCCAGAACTACAGAGGAATGGACCACAGTCTTCCAGG GTCCCTCCTCGTCCACCAACCGGGAGCCTCCACCTTGGCCATCCGTCAGCTATGAATGGCTT TTTAAACAAACCCACGTCCCAGCCTGGGTAACATGGTAAAGCCCCGTCTCTACAAAAAAATC CAAGTTAGCCGGGCATGGTGGTGCGCACCTGTAGTCCCAGCTGCAGTGGGACTGAGGTGGAG GTGGAGGTGGGGGGGGGGGGGAGGGAGGGAGGGATCGCTTGAGCCTGGGAAGTCGAGGCTGC AGTGAGCTGAGATTGCACCACTGCACTCCAGCCTGGGTGACAGAGCAAGACCCTGTCTCAAAAA

FIGURE 256

MSCVLGGVIPLGLLFLVCGSQGYLLPNVTLLEELLSKYQHNESHSRVRRAIPREDKEEILML HNKLRGQVQPQASNMEYMVSAGSGRRGWHRGWGLGHQPALFPSQLCSPASACDGWLRVSSGR GGSRLCSVLFVCFETGSHSATDAGVQWHNRHALKP

Important features:

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 27-31, 41-45

N-myristoylation site.

amino acids 126-132, 140-146

Amidation site.

amino acids 85-89

FIGURE 257

FIGURE 258

 ${\tt MGSGLPLVLLLTLLGSSHGTGPGMTLQLKLKESFLTNSSYESSFLELLEKLCLLLHLPSGTS} \\ {\tt VTLHHARSQHHVVCNT}$

FIGURE 259

FIGURE 260

MIGYYLILFLMWGSSTVFCVLLIFTIAEASFSVENECLVDLCLLRICYKLSGVPNQCRVPLP SDCSK

Important features:

Signal peptide:

amino acids 1-29

FIGURE 261

GAGGATTTGCCACAGCAGCGGATAGAGCAGGAGAGCCCCCGGAGCCCTTGAGACATCCTT GAGAAGAGCCACAGCATAAGAGACTGCCCTGCTTGGTGTTTTTGCAGG**ATG**ATGGTGGCCCTT CGAGGAGCTTCTGCATTGCTGGTTCTGTTCCTTGCAGCTTTTCTGCCCCCCGCCGCAGTGTAC CCAGGACCCAGCCATGGTGCATTACATCTACCAGCGCTTTCGAGTCTTGGAGCAAGGGCTGG AAAAATGTACCCAAGCAACGAGGGCATACATTCAAGAATTCCAAGAGTTCTCAAAAAAATATA TCTGTCATGCTGGGAAGATGTCAGACCTACACAAGTGAGTACAAGAGTGCAGTGGGTAACTT ACGAGTGCATCGTATCAGAGGACAAGACACTGGCAGAAATGTTGCTCCAAGAAGCTGAAGAA GAGAAAAAGATCCGGACTCTGCTGAATGCAAGCTGTGACAACATGCTGATGGGCATAAAGTC TTTGAAAATAGTGAAGAAGATGATGGACACACATGGCTCTTGGATGAAAGATGCTGTCTATA ACTCTCCAAAGGTGTACTTATTAATTGGATCCAGAAACAACACTGTTTGGGAATTTGCAAAC ATACGGGCATTCATGGAGGATAACACCAAGCCAGCTCCCCGGAAGCAAATCCTAACACTTTC CTGGCAGGGAACAGGCCAAGTGATCTACAAAGGTTTTCTATTTTTCATAACCAAGCAACTT CTAATGAGATAATCAAATATAACCTGCAGAAGAGGGACTGTGGAAGATCGAATGCTGCTCCCA GGAGGGGTAGGCCGAGCATTGGTTTACCAGCACTCCCCCTCAACTTACATTGACCTGGCTGT GGATGAGCATGGGCCTCTGGGCCATCCACTCTGGGCCAGGCACCCATAGCCATTTGGTTCTCA CAAAGATTGAGCCGGGCACACTGGGAGTGGAGCATTCATGGGATACCCCATGCAGAAGCCAG GATGCTGAAGCCTCATTCCTCTTGTGTGGGGTTCTCTATGTGGTCTACAGTACTGGGGGCCA GGGCCCTCATCGCATCACCTGCATCTATGATCCACTGGGCACTATCAGTGAGGAGGACTTGC CCAACTTGTTCTTCCCCAAGAGACCAAGAAGTCACTCCATGATCCATTACAACCCCAGAGAT AAGCAGCTCTATGCCTGGAATGAAGGAAACCAGATCATTTACAAACTCCAGACAAAGAGAAA GCTGCCTCTGAAGTAATGCATTACAGCTGTGAGAAAGAGCACTGTGGCTTTGGCAGCTGTTC AGTGTGTAGAAGTGGAAATACGTATGCCTCCTTTCCCAAATGTCACTGCCTTAGGTATCTTC CAAGAGCTTAGATGAGAGCATATCATCAGGAAAGTTTCAACAATGTCCATTACTCCCCCAAA CCTCCTGGCTCTCAAGGATGACCACATTCTGATACAGCCTACTTCAAGCCTTTTGTTTTACT CCCTAATATTCACCACTGGCTTTTCTCTCCCCTGGCCTTTGCTGAAGCTCTTCCCTCTTTTT CAAATGTCTATTGATATTCTCCCATTTTCACTGCCCAACTAAAATACTATTAATATTTCTTT CTTTTCTTTTCTTTTTTTGAGACAAGGTCTCACTATGTTGCCCAGGCTGGTCTCAAACTCC AGAGCTCAAGAGATCCTCCTGCCTCAGCCTCCTAAGTACCTGGGATTACAGGCATGTGCCAC CACACCTGGCTTAAAATACTATTTCTTATTGAGGTTTAACCTCTATTTCCCCTAGCCCTGTC CTTCCACTAAGCTTGGTAGATGTAATAATAAAGTGAAAATATTAACATTTGAATATCGCTTT CCAGGTGTGGAGTGTTTGCACATCATTGAATTCTCGTTTCACCTTTGTGAAACATGCACAAG TCTTTACAGCTGTCATTCTAGAGTTTAGGTGAGTAACACAATTACAAAGTGAAAGATACAGC TAGAAAATACTACAAATCCCATAGTTTTTCCATTGCCCAAGGAAGCATCAAATACGTATGTT TGTTCACCTACTCTTATAGTCAATGCGTTCATCGTTTCAGCCTAAAAATAATAGTCTGTCCC TTTAGCCAGTTTTCATGTCTGCACAAGACCTTTCAATAGGCCTTTCAAATGATAATTCCTCC AGAAAACCAGTCTAAGGGTGAGGACCCCAACTCTAGCCTCCTCTTGTCTTGCTGTCCTCTGT

FIGURE 262

MMVALRGASALLVLFLAAFLPPPQCTQDPAMVHYIYQRFRVLEQGLEKCTQATRAYIQEFQE FSKNISVMLGRCQTYTSEYKSAVGNLALRVERAQREIDYIQYLREADECIVSEDKTLAEMLL QEAEEEKKIRTLLNASCDNMLMGIKSLKIVKKMMDTHGSWMKDAVYNSPKVYLLIGSRNNTV WEFANIRAFMEDNTKPAPRKQILTLSWQGTGQVIYKGFLFFHNQATSNEIIKYNLQKRTVED RMLLPGGVGRALVYQHSPSTYIDLAVDEHGLWAIHSGPGTHSHLVLTKIEPGTLGVEHSWDT PCRSQDAEASFLLCGVLYVVYSTGGQGPHRITCIYDPLGTISEEDLPNLFFPKRPRSHSMIH YNPRDKQLYAWNEGNQIIYKLQTKRKLPLK

FIGURE 263

 ${\tt GGGCGCCGCGTACTCACTAGCTGAGGTGGCAGTGGTTCCACCAAC} {\color{red} {\bf ATG}} {\tt GAGCTCTCGCAGA}$ TGTCGGAGCTCATGGGGCTGTCGGTGTTGCTTGGGCTGCTGGCCCTGATGGCGACGGCGGCG AAATGGATTTCCACCTGACAAATCTTCGGGATCCAAGAAGCAGAAACAATATCAGCGGATTC GGAAGGAGAAGCCTCAACAACACAACTTCACCCACCGCCTCCTGGCTGCAGCTCTGAAGAGC CACAGCGGGAACATATCTTGCATGGACTTTAGCAGCAATGGCAAATACCTGGCTACCTGTGC AGATGATCGCACCATCCGCATCTGGAGCACCAAGGACTTCCTGCAGCGAGAGCACCGCAGCA TGAGAGCCAACGTGGAGCTGGACCACGCCACCCTGGTGCGCTTCAGCCCTGACTGCAGAGCC TTCATCGTCTGGCTGGCCAACGGGGACACCCTCCGTGTCTTCAAGATGACCAAGCGGGAGGA TGGGGGCTACACCTTCACAGCCACCCCAGAGGACTTCCCTAAAAAGCACAAGGCGCCTGTCA TCGACATTGGCATTGCTAACACAGGGAAGTTTATCATGACTGCCTCCAGTGACACCACTGTC ACACGCTGCTGTATCTCCCTGTGGCAGATTTGTAGCCTCGTGTGGCTTCACCCCAGATGTGA AGGTTTGGGAAGTCTGCTTTGGAAAGAAGGGGGGAGTTCCAGGAGGTGGTGCGAGCCTTCGAA CTAAAGGGCCACTCCGCGGCTGTGCACTCGTTTGCTTTCTCCAACGACTCACGGAGGATGGC TTCTGTCTCCAAGGATGGTACATGGAAACTGTGGGACACAGATGTGGAATACAAGAAGAAGC AGGACCCCTACTTGCTGAAGACAGGCCGCTTTGAAGAGGCGGCGGGTGCCGCCGTGCCGC CTGGCCCTCTCCCCCAACGCCCAGGTCTTGGCCTTGGCCAGTGGCAGTAGTATTCATCTCTA CAATACCCGGCGGGGCGAGAAGGAGGAGTGCTTTGAGCGGGTCCATGGCGAGTGTATCGCCA ACTTGTCCTTTGACATCACTGGCCGCTTTCTGGCCTCCTGTGGGGACCGGGCGGTGCGGCTG TTTCACAACACTCCTGGCCACCGAGCCATGGTGGAGGAGATGCAGGGCCACCTGAAGCGGGC CTCCAACGAGAGCACCCGCCAGAGGCTGCAGCAGCTGACCCAGGCCCAAGAGACCCTGA GGATCTGGCCTCCTCATGGCACTGCTGCCATCTTTCCTCCCAGGTGGAAGCCTTTCAGAAGG AGTCTCCTGGTTTTCTTACTGGTGGCCCTGCTTCTTCCCATTGAAACTACTCTTGTCTACTT AGGTCTCTCTTCTTGCTGGCTGTGACTCCTCCCTGACTAGTGGCCAAGGTGCTTTTCTTC CTCCCAGGCCCAGTGGGTGGAATCTGTCCCCACCTGGCACTGAGGAGAATGGTAGAGAGGAG AGGAGAGAGAGAGAATGTGATTTTTGGCCTTGTGGCAGCACATCCTCACACCCAAAGAAG TTTGTAAATGTTCCAGAACAACCTAGAGAACACCTGAGTACTAAGCAGCAGTTTTGCAAGGA TGGGAGACTGGGATAGCTTCCCATCACAGAACTGTGTTCCATCAAAAAGACACTAAGGGATT TCCTTCTGGGCCTCAGTTCTATTTGTAAGATGGAGAATAATCCTCTCTGTGAACTCCTTGCA AAGATGATATGAGGCTAAGAGAATATCAAGTCCCCAGGTCTGGAAGAAAAGTAGAAAAGAGT AGTACTATTGTCCAATGTCATGAAAGTGGTAAAAGTGGGAACCAGTGTGCTTTGAAACCAAA TTAGAAACACATTCCTTGGGAAGGCAAAGTTTTCTGGGACTTGATCATACATTTTATATGGT TGGGACTTCTCTCTGGGAGATGATATCTTGTTTAAGGAGACCTCTTTTCAGTTCATCAAG

FIGURE 264

MELSQMSELMGLSVLLGLLALMATAAVARGWLRAGEERSGRPACQKANGFPPDKSSGSKKQK
QYQRIRKEKPQQHNFTHRLLAAALKSHSGNISCMDFSSNGKYLATCADDRTIRIWSTKDFLQ
REHRSMRANVELDHATLVRFSPDCRAFIVWLANGDTLRVFKMTKREDGGYTFTATPEDFPKK
HKAPVIDIGIANTGKFIMTASSDTTVLIWSLKGQVLSTINTNQMNNTHAAVSPCGRFVASCG
FTPDVKVWEVCFGKKGEFQEVVRAFELKGHSAAVHSFAFSNDSRRMASVSKDGTWKLWDTDV
EYKKKQDPYLLKTGRFEEAAGAAPCRLALSPNAQVLALASGSSIHLYNTRRGEKEECFERVH
GECIANLSFDITGRFLASCGDRAVRLFHNTPGHRAMVEEMQGHLKRASNESTRQRLQQQLTQ
AQETLKSLGALKK

Important features:

Signal peptide:

amino acids 1-25

N-glycosylation site.

amino acids 76-80, 92-96, 231-235, 289-293, 378-382, 421-425

Beta-transducin family Trp-Asp repeat protein.

amino acids 30-47, 105-118, 107-119, 203-216, 205-217, 296-308

FIGURE 265

TGGCCTCCCCAGCTTGCCAGGCACAAGGCTGAGCGGAGGAAGCGAGAGGCATCTAAGCAGG ${\tt CAGTGTTTTGCCTTCACCCCAAGTGACC} {\color{blue} \textbf{ATG}} {\color{blue} \textbf{AGAGGTGCCACGCGAGTCTCAATCATGCTCC}}$ TCCTAGTAACTGTGTCTGACTGTGCTGTGATCACAGGGGCCTGTGAGCGGGATGTCCAGTGT GGGGCAGGCACCTGTGCCATCAGCCTGTGGCTTCGAGGGCTGCGGATGTGCACCCCGCT GGGGCGGAAGGCGAGGAGTGCCACCCCGGCAGCCACAAGGTCCCCTTCTTCAGGAAACGCA AGCACCACACCTGTCCTTGCTTGCCCAACCTGCTGTGCTCCAGGTTCCCGGACGGCAGGTAC $\texttt{CGCTGCTCCATGGACTTGAAGAACATCAATTTT} \underline{\textbf{TAG}} \\ \texttt{GCGCTTGCCTGGTCTCAGGATACCCA}$ CCATCCTTTTCCTGAGCACAGCCTGGATTTTTTTTTTCTGCCATGAAACCCAGCTCCCATGAC TCTCCCAGTCCCTACACTGACTACCCTGATCTCTCTTGTCTAGTACGCACATATGCACACAG GCAGACATACCTCCCATCATGACATGGTCCCCAGGCTGGCCTGAGGATGTCACAGCTTGAGG CTGTGGTGTGAAAGGTGGCCAGCCTGGTTCTCTTCCCTGCTCAGGCTGCCAGAGAGGTGGTA AATGGCAGAAAGGACATTCCCCCTCCCCTCCCCAGGTGACCTGCTCTCTTTCCTGGGCCCTG CCCCTCTCCCCACATGTATCCCTCGGTCTGAATTAGACATTCCTGGGCACAGGCTCTTGGGT GCATTGCTCAGAGTCCCAGGTCCTGGCCTGACCCTCAGGCCCTTCACGTGAGGTCTGTGAGG ACCAATTTGTGGGTAGTTCATCTTCCCTCGATTGGTTAACTCCTTAGTTTCAGACCACAGAC TCAAGATTGGCTCTTCCCAGAGGGCAGCAGACAGTCACCCCAAGGCAGGTGTAGGGAGCCCA GGGAGGCCAATCAGCCCCCTGAAGACTCTGGTCCCAGTCAGCCTGTGGCTTGTGGCCTGTGA CCTGTGACCTTCTGCCAGAATTGTCATGCCTCTGAGGCCCCCTCTTACCACACTTTACCAGT TAACCACTGAAGCCCCCAATTCCCACAGCTTTTCCATTAAAATGCAAATGGTGGTGGTTCAA TCTAATCTGATATTGACATATTAGAAGGCAATTAGGGTGTTTCCTTAAACAACTCCTTTCCA AGGATCAGCCCTGAGAGCAGGTTGGTGACTTTGAGGAGGGCAGTCCTCTGTCCAGATTGGGG TGGGAGCAAGGGACAGGGAGCAGGGCAGGGGCTGAAAGGGGCACTGATTCAGACCAGGGAGG CAACTACACCCAACATGCTGGCTTTAGAATAAAAGCACCAACTGAAAAAA

FIGURE 266

 ${\tt MRGATRVSIMLLLVTVSDCAVITGACERDVQCGAGTCCAISLWLRGLRMCTPLGREGEECHP} \\ {\tt GSHKVPFFRKRKHHTCPCLPNLLCSRFPDGRYRCSMDLKNINF}$

Signal peptide:

amino acids 1-19

Tyrosine kinase phosphorylation site:

amino acids 88-95

N-myristoylation sites:

amino acids 33-39, 35-41, 46-52

FIGURE 267

AGCGCCCGGGCGTCGGGGCGGTAAAAGGCCGGCAGAAGGGAGGCACTTGAGAA**ATG**TCTTTC CTCCAGGACCCAAGTTTCTTCACCATGGGGATGTGGTCCATTGGTGCAGGAGCCCTGGGGGC TGCTGCCTTGGCATTGCTGCTTGCCAACACAGACGTGTTTCTGTCCAAGCCCCAGAAAGCGG CCCTGGAGTACCTGGAGGATATAGACCTGAAAACACTGGAGAAGGAACCAAGGACTTTCAAA GCAAAGGAGCTATGGGAAAAAATGGAGCTGTGATTATGGCCGTGCGGAGGCCAGGCTGTTT CCTCTGTCGAGAGGAAGCTGCGGATCTGTCCTCCCTGAAAAGCATGTTGGACCAGCTGGGCG TCCCCCTCTATGCAGTGGTAAAGGAGCACATCAGGACTGAAGTGAAGGATTTCCAGCCTTAT TTCAAAGGAGAAATCTTCCTGGATGAAAAGAAAAGTTCTATGGTCCACAAAGGCGGAAGAT GATGTTTATGGGATTTATCCGTCTGGGAGTGTGGTACAACTTCTTCCGAGCCTGGAACGGAG GCTTCTCTGGAAACCTGGAAGGAGAAGGCTTCATCCTTGGGGGAGTTTTCGTGGTGGGATCA GGAAAGCAGGGCATTCTTCTTGAGCACCGAGAAAAAGAATTTGGAGACAAAGTAAACCTACT TTCTGTTCTGGAAGCTGCTAAGATGATCAAACCACAGACTTTGGCCTCAGAGAAAAAAA**TGA**T TGTGTGAAACTGCCCAGCTCAGGGATAACCAGGGACATTCACCTGTGTTCATGGGATGTATT GTTTCCACTCGTGTCCCTAAGGAGTGAGAAACCCATTTATACTCTACTCTCAGTATGGATTA TTAATGTATTTTAATATTCTGTTTAGGCCCACTAAGGCAAAATAGCCCCAAAACAAGACTGA CAAAAATCTGAAAAACTAATGAGGATTATTAAGCTAAAACCTGGGAAATAGGAGGCTTAAAA TTGACTGCCAGGCTGGGTGCAGTGGCTCACACCTGTAATCCCAGCACTTTGGGAGGCCAAGG TGAGCAAGTCACTTGAGGTCGGGAGTTCGAGACCAGCCTGAGCAACATGGCGAAACCCCGTC GGAGGCTGAGGCAGGAGATCACTTGAACCTGGGAGGTGGAGGTTGCGGTGAGCTGAGATCA

FIGURE 268

MSFLQDPSFFTMGMWSIGAGALGAAALALLLANTDVFLSKPQKAALEYLEDIDLKTLEKEPR TFKAKELWEKNGAVIMAVRRPGCFLCREEAADLSSLKSMLDQLGVPLYAVVKEHIRTEVKDF QPYFKGEIFLDEKKKFYGPQRRKMMFMGFIRLGVWYNFFRAWNGGFSGNLEGEGFILGGVFV VGSGKQGILLEHREKEFGDKVNLLSVLEAAKMIKPQTLASEKK

FIGURE 269

FIGURE 270

MANPGLGLLLALGLPFLLARWGRAWGQIQTTSANENSTVLPSSTSSSSDGNLRPEAITAIIV VFSLLAALLLAVGLALLVRKLREKRQTEGTYRPSSEEQFSHAAEARAPQDSKETVQGCLPI

FIGURE 271

FIGURE 272

 ${\tt MTFFLSLLLLLVCEAIWRSNSGSNTLENGYFLSRNKENHSQPTQSSLEDSVTPTKAVKTTGK} \\ {\tt GIVKGRNLDSRGLILGAEAWGRGVKKNT}$

FIGURE 273

GCCAGGAATAACTAGAGGGAACA**ATG**GGGTTATTCAGAGGTTTTTGTTTTCCTCTTAGTTCT GTGCCTGCTGCACCAGTCAAATACTTCCTTCATTAAGCTGAATAATAATGGCTTTGAAGATA TTGTCATTGTTATAGATCCTAGTGTGCCAGAAGATGAAAAAATAATTGAACAAATAGAGGAT TGTATCTATATTAATTCCTGAGAATTGGAAGGAAAATCCTCAGTACAAAAGGCCAAAACATG AAAACCATAAACATGCTGATGTTATAGTTGCACCACCTACACTCCCAGGTAGAGATGAACCA TACACCAAGCAGTTCACAGAATGTGGAGAGAAAGGCGAACACTCCACCCTGACCT TCTACTTGGAAAAAACAAAATGAATATGGACCACCAGGCAAACTGTTTGTCCATGAGTGGG CTCACCTCCGGTGGGGAGTGTTTGATGAGTACAATGAAGATCAGCCTTTCTACCGTGCTAAG TCAAAAAAATCGAAGCAACAAGGTGTTCCGCAGGTATCTCTGGTAGAAATAGAGTTTATAA GTGTCAAGGAGGCAGCTGTCTTAGTAGAGCATGCAGAATTGATTCTACAACAAAACTGTATG GAAAAGATTGTCAATTCTTTCCTGATAAAGTACAAACAGAAAAAGCATCCATAATGTTTATG CAAAGTATTGATTCTGTTGTAATTTTGTAACGAAAAAACCCATAATCAAGAAGCTCCAAG CCTACAAAACATAAAGTGCAATTTTAGAAGTACATGGGAGGTGATTAGCAATTCTGAGGATT TTAAAAACACCATACCCATGGTGACACCACCTCCTCCACCTGTCTTCTCATTGCTGAAGATC AGTCAAAGAATTGTGTGCTTAGTTCTTGATAAGTCTGGAAGCATGGGGGGTAAGGACCGCCT AAATCGAATGAATCAAGCAGCAAAACATTTCCTGCTGCAGACTGTTGAAAATGGATCCTGGG TGGGGATGGTTCACTTTGATAGTACTGCCACTATTGTAAATAAGCTAATCCAAATAAAAAGC AGTGATGAAAGAAACACACTCATGGCAGGATTACCTACATATCCTCTGGGAGGAACTTCCATCTGCTCTGGGAATTAAATATGCATTTCAGGTGATTGGAGAGCTACATTCCCAACTCGATGGAT CCGAAGTACTGCTGACTGATGGGGAGGATAACACTGCAAGTTCTTGTATTGATGAAGTG AAACAAAGTGGGGCCATTGTTCATTTTATTGCTTTGGGAAGAGCTGCTGATGAAGCAGTAAT AGAGATGAGCAAGATAACAGGAGGAAGTCATTTTTATGTTTCAGATGAAGCTCAGAACAATG GCCTCATTGATGCTTTTGGGGCTCTTACATCAGGAAATACTGATCTCTCCCAGAAGTCCCTT CAGCTCGAAAGTAAGGGATTAACACTGAATAGTAATGCCTGGATGAACGACACTGTCATAAT TGATAGTACAGTGGGAAAGGACACGTTCTTTCTCATCACATGGAACAGTCTGCCTCCCAGTA TTTCTCTCTGGGATCCCAGTGGAACAATAATGGAAAATTTCACAGTGGATGCAACTTCCAAA ATGGCCTATCTCAGTATTCCAGGAACTGCAAAGGTGGGCACTTGGGCATACAATCTTCAAGC CAAAGCGAACCCAGAAACATTAACTATTACAGTAACTTCTCGAGCAGCAAATTCTTCTGTGC CTCCAATCACAGTGAATGCTAAAATGAATAAGGACGTAAACAGTTTCCCCAGCCCAATGATT GTTTACGCAGAAATTCTACAAGGATATGTACCTGTTCTTGGAGCCAATGTGACTGCTTTCAT TGAATCACAGAATGGACATACAGAAAGTTTTGGAACTTTTTGGATAATGGTGCAGGCGCTGATT AGCTTAAAAGTTCGGGCTCATGGAGGAGCAAACACTGCCAGGCTAAAATTACGGCCTCCACT GAATAGAGCCGCGTACATACCAGGCTGGGTAGTGAACGGGGAAATTGAAGCAAACCCGCCAA GACCTGAAATTGATGAGGATACTCAGACCACCTTGGAGGATTTCAGCCGAACAGCATCCGGA GGTGCATTTGTGGTATCACAAGTCCCAAGCCTTCCCTTGCCTGACCAATACCCACCAAGTCA AATCACAGACCTTGATGCCACAGTTCATGAGGATAAGATTATTCTTACATGGACAGCACCAG GAGATAATTTTGATGTTGGAAAAGTTCAACGTTATATCATAAGAATAAGTGCAAGTATTCTT GATCTAAGAGACAGTTTTGATGATGCTCTTCAAGTAAATACTACTGATCTGTCACCAAAGGA GGCCAACTCCAAGGAAAGCTTTGCATTTAAACCAGAAAATATCTCAGAAGAAAATGCAACCC ACATATTATTGCCATTAAAAGTATAGATAAAAGCAATTTGACATCAAAAGTATCCAACATT GCACAAGTAACTTTGTTTATCCCTCAAGCAAATCCTGATGACATTGATCCTACACCTACTCC TACTCCTACTCCTGATAAAAGTCATAATTCTGGAGTTAATATTTCTACGCTGGTAT TGTCTGTGATTGGGTCTGTTGTAATTGTTAACTTTATTTTAAGTACCACCATT ${f TGA}$ ACCTTAACGAAGAAAAAATCTTCAAGTAGACCTAGAAGAGAGTTTTAAAAAAACAAAACAATGTAAGT

FIGURE 274

MGLFRGFVFLLVLCLLHQSNTSFIKLNNNGFEDIVIVIDPSVPEDEKIIEQIEDMVTTASTY
LFEATEKRFFFKNVSILIPENWKENPQYKRPKHENHKHADVIVAPPTLPGRDEPYTKQFTEC
GEKGEYIHFTPDLLLGKKQNEYGPPGKLFVHEWAHLRWGVFDEYNEDQPFYRAKSKKIEATR
CSAGISGRNRVYKCQGGSCLSRACRIDSTTKLYGKDCQFFPDKVQTEKASIMFMQSIDSVVE
FCNEKTHNQEAPSLQNIKCNFRSTWEVISNSEDFKNTIPMVTPPPPPVFSLLKISQRIVCLV
LDKSGSMGGKDRLNRMNQAAKHFLLQTVENGSWVGMVHFDSTATIVNKLIQIKSSDERNTLM
AGLPTYPLGGTSICSGIKYAFQVIGELHSQLDGSEVLLLTDGEDNTASSCIDEVKQSGAIVH
FIALGRAADEAVIEMSKITGGSHFYVSDEAQNNGLIDAFGALTSGNTDLSQKSLQLESKGLT
LNSNAWMNDTVIIDSTVGKDTFFLITWNSLPPSISLWDPSGTIMENFTVDATSKMAYLSIPG
TAKVGTWAYNLQAKANPETLTITVTSRAANSSVPPITVNAKMNKDVNSFPSPMIVYAEILQG
YVPVLGANVTAFIESQNGHTEVLELLDNGAGADSFKNDGVYSRYFTAYTENGRYSLKVRAHG
GANTARLKLRPPLNRAAYIPGWVVNGEIEANPPRPEIDEDTQTTLEDFSRTASGGAFVVSQV
PSLPLPDQYPPSQITDLDATVHEDKIILTWTAPGDNFDVGKVQRYIIRISASILDLRDSFDD
ALQVNTTDLSPKEANSKESFAFKPENISEENATHIFIAIKSIDKSNLTSKVSNIAQVTLFIP
QANPDDIDPTPTPTPTPTPDKSHNSGVNISTLVLSVIGSVVIVNFILSTTI

Signal peptide:

amino acids 1-21

Putative transmembrane domains:

amino acids 284-300, 617-633

Leucine zipper pattern.

amino acids 469-491, 476-498

N-glycosylation site.

amino acids 20-24, 75-79, 340-344, 504-508, 542-546, 588-592, 628-632, 811-815, 832-836, 837-841, 852-856, 896-900

FIGURE 275

CTCCTTAGGTGGAAACCCTGGGAGTAGAGTACTGACAGCAAAGACCGGGAAAGACCATACGTCCCCG CCCACGGGAAGGAGGGGTACCAGGCCGTCCTTCAGGAGTGGAGGAGCAGCACCGCAACTACGTGA
GCAGCCTGAAGCGGCAGATCGCACAGCTCAAGGAGGAGGAGGAGGAGGAGCAGCTCAGGAA
TGGGCAGTACCAAGCCAGCGATGCTGGCCTGGGTCTGGACAGGAGAGCCCCCCAGAGAAAACCCAG
GCCGACCTCCTGGCCTTCCTGCACTCGCAGTGGACAAGGCAGAGGTGAATGCTGGCGTCAAGCTGG
CCACAGAGTATGCAGCAGTGCCTTTCGATAGCTTTACTCTACAGAAGGTGTACCAGCTGGAGACTGG
CCTTACCCGCCACCCCGAGGAGAAGCCTGTGAGGAAGGACAAGCGGGATGAGTTTGGTGGAAGCCATT
GAATCAGCCTTGGAGACCCTTGAACAATCCTGCAGAGAACAGCCCCAATCACCGTCCTTACCACGCCT
CTGATTTCATAGAAGGGATCTACCGAACAAAGGGACAAAGGGACATTCAGCCCCATCATGAAAGTG
AAAATGAAAAGCTCAACATGGCCAACACCGCTTTATCAATGTTATCCGCCCCTCTAGCAAAAAAGGGACCACCA AAAAATGAAAAGCTCAACATGGCCAACACGCTTATCAATGTTATCGTGCCTCTAGCAAAAAGGGTGG ACAAGTTCCGGCAGTTCATGCAGAATTTCAGGGAGATGTGCATTGAGCAGGATGGGAGAGTCCATCT CACTGTTGTTTACTTTGGGAAAGAAAATAAATGAAGTCAAAGGAATACTTGAAAACACTTCCAAA CTTCGCAAACAGAAACAGAAGCAAGTAGCAAAAAAAACA**TGA**ACTCCCAGAGAAGGATTGTGGGAGA CACTTTTCTTTCCTTTTGCAATTACTGAAAGTGGCTGCAACAGAGAAAAGACTTCCATAAAGGACG ACAAAAGAATTGGACTGATGGGTCAGAGATGAGAAAGCCTCCGATTTCTCTGTTGGGCTTTTTAC AACAGAAATCAAAATCTCCGCTTTGCCTGCAAAAGTAACCCAGTTGCACCCTGTGAAGTGTCTGACA AAGGCAGAATGCTTGTGAGATTATAAGCCTAATGGTGTGGAGGTTTTGATGGTGTTTTACAATACACT AGGCCATGAATGGAAGGTGGTATTGCACAGCTAATAAAATATGATTTGTGGATATGAA

FIGURE 276

MMMVRRGLLAWISRVVVLLVLLCCAISVLYMLACTPKGDEEQLALPRANSPTGKEGYQAVLQ
EWEEQHRNYVSSLKRQIAQLKEELQERSEQLRNGQYQASDAAGLGLDRSPPEKTQADLLAFL
HSQVDKAEVNAGVKLATEYAAVPFDSFTLQKVYQLETGLTRHPEEKPVRKDKRDELVEAIES
ALETLNNPAENSPNHRPYTASDFIEGIYRTERDKGTLYELTFKGDHKHEFKRLILFRPFSPI
MKVKNEKLNMANTLINVIVPLAKRVDKFRQFMQNFREMCIEQDGRVHLTVVYFGKEEINEVK
GILENTSKAANFRNFTFIQLNGEFSRGKGLDVGARFWKGSNVLLFFCDVDIYFTSEFLNTCR
LNTQPGKKVFYPVLFSQYNPGIIYGHHDAVPPLEQQLVIKKETGFWRDFGFGMTCQYRSDFI
NIGGFDLDIKGWGGEDVHLYRKYLHSNLIVVRTPVRGLFHLWHEKRCMDELTPEQYKMCMQS
KAMNEASHGQLGMLVFRHEIEAHLRKQKQKTSSKKT

FIGURE 277

GAAAGA**TG**TTGTGGCTGCTCTTTTTTCTGGTGACTGCCATTCATGCTGAACTCTGTCAACC AGGTGCAGAAAATGCTTTTAAAGTGAGACTTAGTATCAGAACAGCTCTGGGAGATAAAGCAT ATGCCTGGGATACCAATGAAGAATACCTCTTCAAAGCGATGGTAGCTTTCTCCATGAGAAAA GTTCCCAACAGAAACCAACAGAAATTTCCCATGTCCTACTTTGCAATGTAACCCAGAGGGT ATCATTCTGGTTTGTGGTTACAGACCCTTCAAAAAATCACACCCTTCCTGCTGTTGAGGTGC CTGGAATTTTTAAAAATCCCTTCCACACTTGCACCACCCATGGACCCATCTGTGCCCATCTG GATTATTATATTTGGTGTGATATTTTGCATCATCATAGTTGCAATTGCACTACTGATTTTAT CAGGGATCTGGCAACGTAGAAGAAGAACAAAGAACCATCTGAAGTGGATGACGCTGAAGAT AAGTGTGAAAACATGATCACAATTGAAAATGGCATCCCCTCTGATCCCCTGGACATGAAGGG GGGCATATTAATGATGCCTTCA**TGA**CAGAGGATGAGAGGCTCACCCCTCTCTGAAGGGCTGT TGTTCTGCTTCCTCAAGAAATTAAACATTTGTTTCTGTGTGACTGCTGAGCATCCTGAAATA CCAAGAGCAGATCATATTTTGTTTCACCATTCTTCTTTTGTAATAAATTTTGAATGTGCT TGAAAGTGAAAAGCAATCAATTATACCCACCAACACCACTGAAATCATAAGCTATTCACGAC TCAAAATATTCTAAAATATTTTTCTGACAGTATAGTGTATAAATGTGGTCATGTGGTATTTG TAGTTATTGATTTAAGCATTTTTAGAAATAAGATCAGGCATATGTATATATTTTCACACTTC AAAGACCTAAGGAAAAATAAATTTTCCAGTGGAGAATACATATAATATGGTGTAGAAATCAT TGAAAATGGATCCTTTTTGACGATCACTTATATCACTCTGTATATGACTAAGTAAACAAAAG TGAGAAGTAATTATTGTAAATGGATGGATAAAAATGGAATTACTCATATACAGGGTGGAATT TTATCCTGTTATCACACCAACAGTTGATTATATTTTTCTGAATATCAGCCCCTAATAGGAC AATTCTATTTGTTGACCATTTCTACAATTTGTAAAAGTCCAATCTGTGCTAACTTAATAAAG

FIGURE 278

MLWLLFFLVTAIHAELCQPGAENAFKVRLSIRTALGDKAYAWDTNEEYLFKAMVÁFSMRKVP NREATEISHVLLCNVTQRVSFWFVVTDPSKNHTLPAVEVQSAIRMNKNRINNAFFLNDQTLE FLKIPSTLAPPMDPSVPIWIIIFGVIFCIIIVAIALLILSGIWQRRRKNKEPSEVDDAEDKC ENMITIENGIPSDPLDMKGGILMMPS

FIGURE 279

AACTCAAACTCCTCTCTGGGAAAACGCGGTGCTTGCTCCTCCCGGAGTGGCCTTGGCAGG $\tt GTGTTGGAGCCCTCGGTCTGCCCCGTCCGGTCTCTGGGGCCAAGGCTGGGTTTCCCTC \\ \underline{\textbf{ATG}} \\ \texttt{T}$ ATGGCAAGAGCTCTACTCGTGCGGTGCTTCTTCTCCTTGGCATACAGCTCACAGCTCTTTGG CCTATAGCAGCTGTGGAAATTTATACCTCCCGGGTGCTGGAGGCTGTTAATGGGACAGATGC TCGGTTAAAATGCACTTTCTCCAGCTTTGCCCCTGTGGGTGATGCTCTAACAGTGACCTGGA ATTTTCGTCCTCTAGACGGGGGACCTGAGCAGTTTGTATTCTACTACCACATAGATCCCTTC CAACCCATGAGTGGGCGGTTTAAGGACCGGGTGTCTTGGGATGGGAATCCTGAGCGGTACGA TGCCTCCATCCTTCTCTGGAAACTGCAGTTCGACGACAATGGGACATACACCTGCCAGGTGA AGAACCCACCTGATGTTGATGGGGTGATAGGGGAGATCCGGCTCAGCGTCGTGCACACTGTA CGCTTCTCTGAGATCCACTTCCTGGCTCTGGCCATTGGCTCTGCCTGTGCACTGATGATCAT AATAGTAATTGTAGTGGTCCTCTTCCAGCATTACCGGAAAAAGCGATGGGCCGAAAGAGCTC ATAAAGTGGTGGAGATAAAATCAAAAGAAGAGGGAAAGGCTCAACCAAGAGAAAAAGGTCTCT GTTTATTTAGAAGACACAGAC**TAA**CAATTTTAGATGGAAGCTGAGATGATTTCCAAGAACAA GAACCCTAGTATTTCTTGAAGTTAATGGAAACTTTTCTTTGGCTTTTCCAGTTGTGACCCGT TTTCCAACCAGTTCTGCAGCATATTAGATTCTAGACAAGCAACACCCCTCTGGAGCCAGCAC AGTGCTCCTCCATATCACCAGTCATACACAGCCTCATTATTAAGGTCTTATTTAATTTCAGA GTGTAAATTTTTTCAAGTGCTCATTAGGTTTTATAAACAAGAAGCTACATTTTTTGCCCTTAA GACACTACTTACAGTGTTATGACTTGTATACACATATATTGGTATCAAAGGGGATAAAAGCC AATTTGTCTGTTACATTTCCTTTCACGTATTTCTTTTAGCAGCACTTCTGCTACTAAAGTTA ATGTGTTTACTCTTTCCTTCCCACATTCTCAATTAAAAGGTGAGCTAAGCCTCCTCGGTG TTTCTGATTAACAGTAAATCCTAAATTCAAACTGTTAAATGACATTTTTATTTTTATGTCTC TTTGTCG

FIGURE 280

MYGKSSTRAVLLLLGIQLTALWPIAAVEIYTSRVLEAVNGTDARLKCTFSSFAPVGDALTVT WNFRPLDGGPEQFVFYYHIDPFQPMSGRFKDRVSWDGNPERYDASILLWKLQFDDNGTYTCQ VKNPPDVDGVIGEIRLSVVHTVRFSEIHFLALAIGSACALMIIIVIVVVLFQHYRKKRWAER AHKVVEIKSKEEERLNQEKKVSVYLEDTD

FIGURE 281

FIGURE 282

MKFLAVLVLLGVSIFLVSAQNPTTAAPADTYPATGPADDEAPDAETTAAATTATTAAPTTAT TAASTTARKDIPVLPKWVGDLPNGRVCP

FIGURE 283

GGACTCTGAAGGTCCCAAGCAGCTGCTGAGGCCCCCAAGGAAGTGGTTCCAACCTTGGACCC
CTAGGGGTCTGGATTTGCTGGTTAACAAGATAACCTGAGGGCAGGCCCCATAGGGGAATGC
TACCTCCTGCCCTTCCACCTGCCCTGGTGTTCACGGTGGCCTGGTCCCTCCTTGCCGAGAGA
GTGTCCTGGGTCAGGGACGCAGAGGACGCTCACAGACTCCAGCCCTTTGTTACCGAGAGAC
ACTTGGCAAGGTCCAGCGATGGTCCGGAGTCCACACACAGACTGGCGGCAGGGCAGGAGGG
GACAGTTCTGTTGTGCTTGGTTGGACAGTAAGAGGGTCTTGGCCAGTCCAGGGTGGGGGGG
GCAAACTCCATAAAGAACCAGAGGGTCTGGGCCCCGGCCACAGAGTCATCTGCCCAGCTCCT
CTGCTGCTGGCCAGTGGGAGTGGCACGAGGTGGGGGCAGAGTCATCTTGCCCAGCTCCT
CTGCTGCTGCGGGCCATGGTCCCTGTCTAGGGCAGCAATTCTCAACCTTCTTGCTCTCAGGA
ATTTGCCTGCGGGCCATGGTCCCTGTCTAGGGCAGCAATTCTCAACCTTCTTGCTCTCAGGA
CCCCAAAGAGCTTTCATTGTATCTATTGATTTTTACCACATTAGCAATTAAAACTGAGAAAT
GGGCCGGGCACGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGTGGAT
CACCTGAGATCAGGAGTTCAAGACCAGCCTGGCCAACATGGTGAAACCTTGTCTACTAAAAA
TACAAAAAATTAGCCAGGCACAGTGGTGCACTGGTAGTCCCAGTTACTCGGGAGGCTGAG
GCAGGAAAATCGCTTGAACCCAGGAGGCGGACGTTGCGGTGAGCCGAGATCGCCCCCTGAT
TCCAGCCTGGGCGACAAGAGTGAGACTCCATCTCACACA

FIGURE 284

MLPPALPPALVFTVAWSLLAERVSWVRDAEDAHRLQPFVTERTLGKVQRWSGVHTQTGGRAG GGQFCCAWLDSKRVLASPGWGAANSIKNQRVWAPATESSAQLLCCWPVGVARGGALCQ

FIGURE 285

FIGURE 286

MPVPALCLLWALAMVTRPASAAPMGGPELAQHEELTLLFHGTLQLGQALNGVYRTTEGRLTK ARNSLGLYGRTIELLGQEVSRGRDAAQELRASLLETQMEEDILQLQAEATAEVLGEVAQAQK VLRDSVQRLEVQLRSAWLGPAYREFEVLKAHADKQSHILWALTGHVQRQRREMVAQQHRLRQ IQERLHTAALPA

FIGURE 287

GGCAAC**ATG**GCTCAGCAGGCTTGCCCCAGAGCCATGGCAAAGAATGGACTTGTAATTTGCAT CCTGGTGATCACCTTACTCCTGGACCAGACCACCAGCCACACATCCAGATTAAAAGCCAGGA AGCACAGCAAACGTCGAGTGAGAGACAAGGATGGAGATCTGAAGACTCAAATTGAAAAGCTC TAAAGTTCACAAGAAATGCTACCTTGCTTCAGAAGGTTTGAAGCATTTCCATGAGGCCAATG AAGACTGCATTTCCAAAGGAGGAATCCTGGTTATCCCCAGGAACTCCGACGAAATCAACGCC CTCCAAGACTATGGTAAAAGGAGCCTGCCAGGTGTCAATGACTTTTGGCTGGGCATCAATGA ACCGTGCACAGCCTAACGGTGGCAAGCGAGAAAACTGTGTCCTGTTCTCCCAATCAGCTCAG GGCAAGTGGAGTGATGAGGCCTGTCGCAGCAGCAAGAGATACATATGCGAGTTCACCATCCC TAAATAGGTCTTTCTCCAATGTGTCCTCCAAGCAAGATTCATCATAACTTATAGGTTCATGA TCTCTAAGATCAAGTAAAAATCATAATTTTTACTTATTAAAAAATTGCAACACAAGATCAAT GTCCATAGCAATATGATAGCATCAGCCAATTTTGCTAACACATTTCTTTGGGATTTTGCCCT TCCTGGGGTATAGGGGATCAGAAATATTGATCCATGTGCACGCAGATAAAATGGCTTCTGCT TTCCCTTGGAAGTTTAGCGTATGTTTGACTAACAAAAATTCCCTACATCAGAGACTCTAGGT GCTATATAATCCAAAAACTTTTCAGCCTGTTGCTCATTCTGTCCCATGCTGGCAATAATACC TTGTCAGCCCATTACCCTTATTTTGAATTGCTCCATCTCCTGGTGGGACTTGTATCTTGTCT TACCCTTTTTTTGGAAGTTTCCAGCCGCAATTTGAAATGAAATGACAAGGTGTATATTTGAT CAATTTTCATTCCCACCATTGCATTACAACCTCTAACTTAAATGGGTAACCCTAAGGCATAT AGCATCCTTACTCTCACCTTTTATGAGATTGAGAGTGGACTTACATTTCCTTTTTTACATTT TCGTATATTTATTTTTTTTAGCCATCATTATATGTTTAAGTCTATTATGGGCAACCAATCTT TGGAAGCTGAAAACTGAATTTAAAGAATGCTATCTTGGAAAATTGCATACGTCTGTGCAATT TTTTATTCTGCCTAGTGCTATTCTGCTTGTTTAACTAGATTGTACAAAATAACTTCATTGCT TAATATCAAATTACAAAGTTTAGACTTGGAGGGAAATGGGCTTTTTAGAAGCAAACAATTTT AAATATATTTTGTTCTTCAAATAAATAGTGTTTTAAACATTGAATGTGTTTTTGTGAACAATAT CCCACTTTGCAAACTTTAACTACACATGCTTGGAATTAAGTTTTAGCTGTTTTCATTGCTCA

FIGURE 288

MAQQACPRAMAKNGLVICILVITLLLDQTTSHTSRLKARKHSKRRVRDKDGDLKTQIEKLWT EVNALKEIQALQTVCLRGTKVHKKCYLASEGLKHFHEANEDCISKGGILVIPRNSDEINALQ DYGKRSLPGVNDFWLGINDMVTEGKFVDVNGIAISFLNWDRAQPNGGKRENCVLFSQSAQGK WSDEACRSSKRYICEFTIPK

FIGURE 289

FIGURE 290

MKLAALLGLCVALSCSSAAAFLVGSAKPVAQPVAALESAAEAGAGTLANPLGTLNPLKLLLS SLGIPVNHLIEGSQKCVAELGPQAVGAVKALKALLGALTVFG

FIGURE 291

TGAAGGACTTTTCCAGGACCCAAGGCCACACTGGAAGTCTTGCAGCTGAAGGGAGGCACT $\verb| CCTTGGCCTCCGCAGCCGATCAC| \textbf{ATG} | \texttt{AAGGTGGTGCCAAGTCTCCTGCTCCTCCTG}|$ GCACAGGTGTGGCTGGTACCCGGCTTGGCCCCCAGTCCTCAGTCGCCAGAGACCCCAGCCCC TCAGAACCAGACCAGGGTAGTGCAGGCTCCCAGGGAGGAAGAGGAAGATGAGCAGGAGG CCAGCGAGGAGAGGCCGGTGAGGAAGAGAAGCCTGGCTGATGGCCAGCAGCAGCAGCTT GCCAAGGAGACTTCAAACTTCGGATTCAGCCTGCTGCGAAAGATCTCCATGAGGCACGATGG ${\tt CAACATGGTCTTCTCTCCATTTGGCATGTCCTTGGCCATGACAGGCTTGATGCTGGGGGCCA}$ CAGGGCCGACTGAAACCCAGATCAAGAGAGGGGCTCCACTTGCAGGCCCTGAAGCCCACCAAG CCCGGGCTCCTGCCTTTTAAGGGACTCAGAGAGACCCTCTCCCGCAACCTGGAACT GGGCCTCTCACAGGGGAGTTTTGCCTTCATCCACAAGGATTTTGATGTCAAAGAGACTTTCT TCAATTTATCCAAGAGGTATTTTGATACAGAGTGCGTGCCTATGAATTTTCGCAATGCCTCA CAGGCCAAAAGGCTCATGAATCATTACATTAACAAAGAGACTCGGGGGGAAAATTCCCAAACT GTTTGATGAGATTAATCCTGAAACCAAATTAATTCTTGTGGATTACATCTTGTTCAAAGGGA AATGGTTGACCCCATTTGACCCTGTCTTCACCGAAGTCGACACTTTCCACCTGGACAAGTAC AAGACCATTAAGGTGCCCATGATGTACGGTGCAGGCAAGTTTGCCTCCACCTTTGACAAGAA TTTTCGTTGTCATGTCCTCAAACTGCCCTACCAAGGAAATGCCACCATGCTGGTGGTCCTCA TGGAGAAAATGGGTGACCACCTCGCCCTTGAAGACTACCTGACCACAGACTTGGTGGAGACA GAAGTATGAGATGCATGAGCTGCTTAGGCAGATGGGAATCAGAAGAATCTTCTCACCCTTTG CTGACCTTAGTGAACTCTCAGCTACTGGAAGAAATCTCCAAGTATCCAGGGTTTTACGAAGA ACAGTGATTGAAGTTGAAAAGGGGCACTGAGGCAGTGGCAGGAATCTTGTCAGAAATTAC TGCTTATTCCATGCCTCCTGTCATCAAAGTGGACCGGCCATTTCATTTCATGATCTATGAAG ${\tt AAACCTCTGGAATGCTTCTGTTTCTGGGCAGGGTGGTGAATCCGACTCTCCTA}$ TACCAGCAATGGATGGCAGGGGAGAGTGTTCCTTTTGTTCTTAACTAGTTTAGGGTGTTCTC AAATAAATACAGTAGTCCCCACTTATCTGAGGGGGATACATTCAAAGACCCCCAGCAGATGC AAAGTTTAATTTATAAATTAGGCACAGTAAGAGATTAACAATAATAACAACATTAAGTAAAA TGAGTTACTTGAACGCAAGCACTGCAATACCATAACAGTCAAACTGATTATAGAGAAGGCTA CTAAGTGACTCATGGGCGAGGAGCATAGACAGTGTGGAGACATTGGGCAAGGGGAGAATTCA CATCCTGGGTGGGACAGAGCAGGACGATGCAAGATTCCATCCCACTACTCAGAATGGCATGC TGCTTAAGACTTTTAGATTGTTTATTTCTGGAATTTTTCATTTAATGTTTTTTGGACCATGGT TGACCATGGTTAACTGAGACTGCAGAAAGCAAAACCATGGATAAGGGAGGACTACTACAAAA

FIGURE 292

MKVVPSLLLSVLLAQVWLVPGLAPSPQSPETPAPQNQTSRVVQAPREEEEDEQEASEEKAGE
EEKAWLMASRQQLAKETSNFGFSLLRKISMRHDGNMVFSPFGMSLAMTGLMLGATGPTETQI
KRGLHLQALKPTKPGLLPSLFKGLRETLSRNLELGLSQGSFAFIHKDFDVKETFFNLSKRYF
DTECVPMNFRNASQAKRLMNHYINKETRGKIPKLFDEINPETKLILVDYILFKGKWLTPFDP
VFTEVDTFHLDKYKTIKVPMMYGAGKFASTFDKNFRCHVLKLPYQGNATMLVVLMEKMGDHL
ALEDYLTTDLVETWLRNMKTRNMEVFFPKFKLDQKYEMHELLRQMGIRRIFSPFADLSELSA
TGRNLQVSRVLRRTVIEVDERGTEAVAGILSEITAYSMPPVIKVDRPFHFMIYEETSGMLLF
LGRVVNPTLL

FIGURE 293

FIGURE 294

MRRLLLVTSLVVVLLWEAGAVPAPKVPIKMQVKHWPSEQDPEKAWGARVVEPPEKDDQLVVL FPVQKPKLLTTEEKPRGQGRGPILPGTKAWMETEDTLGRVLSPEPDHDSLYHPPPEEDQGEE RPRLWVMPNHQVLLGPEEDQDHIYHPQ

FIGURE 295

TACCCAAGGAAAGTGCAGCTGAGACTCAGACAAGATTACA**ATG**AACCAACTCAGCTTCCTGC TGGACCTGTTCTTCGTCTCCATCTCTGCCCAGAAGCTGCAAGGAAATCAAAGACGAATGTCC TAGTGCATTTGATGGCCTGTATTTTCTCCGCACTGAGAATGGTGTTATCTACCAGACCTTCT GTGACATGACCTCTGGGGGTGGCGGCTGGACCCTGGTGGCCAGCGTGCATGAGAATGACATG CGTGGGAAGTGCACGGTGGCCGATCGCTGGTCCAGTCAGCAGGGCAAAGCAGACTACCC AGAGGGGGACGGCAACTACAACACCTTTGGATCTGCAGAGGCGGCCACGAGCG ATGACTACAAGAACCCTGGCTACTACGACATCCAGGCCAAGGACCTGGGCATCTGGCACGTG CCCAATAAGTCCCCCATGCAGCACTGGAGAAACAGCTCCCTGCTGAGGTACCGCACGGACAC TGGCTTCCTCCAGACACTGGGACATAATCTGTTTGGCATCTACCAGAAATATCCAGTGAAAT ATGGAGAAGGAAAGTGTTGGACTGACAACGGCCCGGTGATCCCTGTGGTCTATGATTTTGGC GACGCCCAGAAAACAGCATCTTATTACTCACCCTATGGCCAGCGGGAATTCACTGCGGGATT TGTTCAGTTCAGGGTATTTAATAACGAGAGCAGCCAACGCCTTGTGTGCTGGAATGAGGG TCACCGGATGTAACACTGAGCATCACTGCATTGGTGGAGGAGGATACTTTCCAGAGGCCAGT CCCCAGCAGTGTGGAGATTTTTCTGGTTTTGATTGGAGTGGATATGGAACTCATGTTGGTTA CAGCAGCAGCCGTGAGATAACTGAGGCAGCTGTGCTTCTATTCTATCGT**TGA**GAGTTTTGTG GGAGGGAACCCAGACCTCTCCCCAACCATGAGATCCCAAGGATGGAGAACAACTTACCCA GTAGCTAGAATGTTAATGGCAGAAGAGAAAACAATAAATCATATTGACTCAAGAAAAAA

FIGURE 296

MNQLSFLLFLIATTRGWSTDEANTYFKEWTCSSSPSLPRSCKEIKDECPSAFDGLYFLRTEN GVIYQTFCDMTSGGGGWTLVASVHENDMRGKCTVGDRWSSQQGSKADYPEGDGNWANYNTFG SAEAATSDDYKNPGYYDIQAKDLGIWHVPNKSPMQHWRNSSLLRYRTDTGFLQTLGHNLFGI YQKYPVKYGEGKCWTDNGPVIPVVYDFGDAQKTASYYSPYGQREFTAGFVQFRVFNNERAAN ALCAGMRVTGCNTEHHCIGGGGYFPEASPQQCGDFSGFDWSGYGTHVGYSSSREITEAAVLL FYR

FIGURE 297

GCGGAGCCGGCCGCCGCAGAGGAGCCGCTCTCGCCGCCGCCACCTCGGCTGGGAGCC ${\tt CACGAGGCTGCCGCATCCTGCCCTCGGAACA} {\color{red} {\bf ATG}} {\tt GGACTCGGCGCGCGAGGTGCTTGGGCCG}$ CGCTGCTCCTGGGGACGCTGCAGGTGCTAGCGCTGCTGGGGGCCCCCATGAAAGCGCAGCC AGAGACTCTCCAACATGTGCCTTCTGACCATACAAATGAAACTTCCAACAGTACTGTGAAAC CACCAACTTCAGTTGCCTCAGACTCCAGTAATACAACGGTCACCACCATGAAACCTACAGCG GCATCTAATACAACACCAGGGATGGTCTCAACAAATATGACTTCTACCACCTTAAAGTC TACACCCAAAACAACAAGTGTTTCACAGAACACATCTCAGATATCAACATCCACAATGACCG TAACCCACAATAGTTCAGTGACATCTGCTGCTTCATCAGTAACAATCACAACAACTATGCAT TCTGAAGCAAAGAAAGGATCAAAATTTGATACTGGGAGCTTTGTTGGTGGTATTGTATTAAC GCTGGGAGTTTTATCTATTCTTTACATTGGATGCAAAATGTATTACTCAAGAAGAGGCATTC GGTATCGAACCATAGATGAACATGATGCCATCATT**TAA**GGAAATCCATGGACCAAGGATGGA ATACAGATTGATGCTGCCCTATCAATTAATTTTGGTTTATTAATAGTTTAAAAACAATATTCT CTTTTTGAAAATAGTATAAACAGGCCATGCATATAATGTACAGTGTATTACGTAAATATGTA AAGATTCTTCAAGGTAACAAGGGTTTGGGTTTTGAAATAAACATCTGGATCTTATAGACCGT GGGGTGGGGGCATTGGTCACATATGACCAGTAATTGAAAGACGTCATCACTGAAAGACAGAA TGCCATCTGGGCATACAAATAAGAAGTTTGTCACAGCACTCAGGATTTTGGGTATCTTTTGT AGCTCACATAAAGAACTTCAGTGCTTTTTCAGAGCTGGATATATCTTAATTACTAATGCCACA CAGAAATTATACAATCAAACTAGATCTGAAGCATAATTTAAGAAAAACATCAACATTTTTTG TGCTTTAAACTGTAGTAGTTGGTCTAGAAACAAAATACTCC

FIGURE 298

MGLGARGAWAALLLGTLQVLALLGAAHESAAMAASANIENSGLPHNSSANSTETLQHVPSDH TNETSNSTVKPPTSVASDSSNTTVTTMKPTAASNTTTPGMVSTNMTSTTLKSTPKTTSVSQN TSQISTSTMTVTHNSSVTSAASSVTITTTMHSEAKKGSKFDTGSFVGGIVLTLGVLSILYIG CKMYYSRRGIRYRTIDEHDAII

FIGURE 299

CAGCCGGGTCCCAAGCCTGTGCCTGAGCCTGAGCCTGAGCCCGAGCCGGGAGCCGG CTTCTTCGGCTTGGCTCAGCCTGTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCT ATAAAGAAAATTCTGGGCATATTTATAATAAGAACATATCTCAGAAAGATTGTGATTGCCTT CATGTTGTGGAGCCCATGCCTGTGCGGGGGCCTGATGTAGAAGCATACTGTCTACGCTGTGA ATGCAAATATGAAGAAGAAGCTCTGTCACAATCAAGGTTACCATTATAATTTATCTCTCCA TTTTGGGCCTTCTACTTCTGTACATGGTATATCTTACTCTGGTTGAGCCCATACTGAAGAGG CGCCTCTTTGGACATGCACAGTTGATACAGAGTGATGATGATATTGGGGGATCACCAGCCTTT TGCAAATGCACACGATGTGCTAGCCCGCTCCCGCAGTCGAGCCAACGTGCTGAACAAGGTAG ${\tt CATGTTGTCCTCAGC} \underline{{\tt TAA}} {\tt TTGGGGAATTGAATTCAAGGTGACTAGAAAGAAACAGGCAGACAA}$ ATAATAGAGACATTTTTAAAAGCACACAGCTCAAAGTCAGCCAATAAGTCTTTTCCTATTTG TGACTTTTACTAATAAAATAAATCTGCCTGTAAATTATCTTGAAGTCCTTTACCTGGAACA AGCACTCTCTTTTCACCACATAGTTTTAACTTGACTTTCAAGATAATTTTCAGGGTTTTTG AACAACTTTTTCAAGTCACTTTACTAAACAAACTTTTGTAAATAGACCTTACCTTCTATTT TCGAGTTTCATTTATATTTTGCAGTGTAGCCAGCCTCATCAAAGAGCTGACTTACTCATTTG ACTTTTGCACTGACTGTATTATCTGGGTATCTGCTGTGTCTGCACTTCATGGTAAACGGGAT CTAAAATGCCTGGTGGCTTTTCACAAAAAGCAGATTTTCTTCATGTACTGTGATGTCTGATG CAATGCATCCTAGAACAAACTGGCCATTTGCTAGTTTACTCTAAAGACTAAACATAGTCTTG GTGTGTGTGTCTTACTCATCTTCTAGTACCTTTAAGGACAAATCCTAAGGACTTGGACACT TGCAATAAAGAAATTTTATTTTAAACCCAAGCCTCCCTGGATTGATAATATATACACATTTG TCAGCATTTCCGGTCGTGAGAGGCAGCTGTTTGAGCTCCAATATGTGCAGCTTTGAACT AGGGCTGGGGTTGTGGGTGCCTCTTCTGAAAGGTCTAACCATTATTGGATAACTGGCTTTTT TCTTCCTATGTCCTCTTTGGAATGTAACAATAAAAATAATTTTTTGAAACATCAA

FIGURE 300

MATLWGGLLRLGSLLSLSCLALSVLLLAQLSDAAKNFEDVRCKCICPPYKENSGHIYNKNIS QKDCDCLHVVEPMPVRGPDVEAYCLRCECKYEERSSVTIKVTIIIYLSILGLLLLYMVYLTL VEPILKRRLFGHAQLIQSDDDIGDHQPFANAHDVLARSRSRANVLNKVEYAQQRWKLQVQEQ RKSVFDRHVVLS

FIGURE 301

FIGURE 302

MAYSTVQRVALASGLVLALSLLLPKAFLSRGKRQEPPPTPEGKLGRFPPMMHHHQAPSDGQT PGARFQRSHLAEAFAKAKGSGGGAGGGGSGRGLMGQIIPIYGFGIFLYILYILFKVSRIILI ILHQ

FIGURE 303

 $\tt CGGCTCGAGTGCAGCTGTGGGGAGATTTCAGTGCATTGCCTCCCCTGGGTGCTCTTCATCTT$ $\mathsf{GGATTTGAAAGTTGAGAGCAGC} \underline{\mathbf{ATG}} \mathtt{TTTTGCCCACTGAAACTCATCCTGCCAGTGTTAC}$ TGGATTATTCCTTGGGCCTGAATGACTTGAATGTTTCCCCGCCTGAGCTAACAGTCCATGTG GGTGATTCAGCTCTGATGGGATGTTTTTCCAGAGCACAGAAGACAAATGTATATTCAAGAT AGACTGGACTCTGTCACCAGGAGAGCACGCCAAGGACGAATATGTGCTATACTATTACTCCA ATCTCAGTGTGCCTATTGGGCGCTTCCAGAACCGCGTACACTTGATGGGGGACATCTTATGC AATGATGGCTCTCCTGCTCCAAGATGTGCAAGAGGCTGACCAGGGAACCTATATCTGTGA AATCCGCCTCAAAGGGGAGAGCCAGGTGTTCAAGAAGGCGGTGGTACTGCATGTGCTTCCAG AGCACAGAAGTGAAACACGTGACCAAGGTAGAATGGATATTTTCAGGACGGCGCGCAAAGGA GGAGATTGTATTTCGTTACTACCACAAACTCAGGATGTCTGTGGAGTACTCCCAGAGCTGGG GCCACTTCCAGAATCGTGTGAACCTGGTGGGGGACATTTTCCGCAATGACGGTTCCATCATG CTTCAAGGAGTGAGGGAGTCAGATGGAGGAAACTACACCTGCAGTATCCACCTAGGGAACCT GGTGTTCAAGAAAACCATTGTGCTGCATGTCAGCCCGGAAGAGCCCTCGAACACTGGTGACCC CGGCAGCCCTGAGGCCTCTGGTCTTGGGTGGTAATCAGTTGGTGATCATTGTGGGAATTGTC TGTGCCACAATCCTGCTGCTCCCTGTTCTGATATTGATCGTGAAGAAGACCTGTGGAAATAA GAGTTCAGTGAATTCTACAGTCTTGGTGAAGAACACGAAGAAGACTAATCCAGAGATAAAAG AAAAACCCTGCCATTTTGAAAGATGTGAAGGGGAGAAACACATTTACTCCCCAATAATTGTA CGGGAGGTGATCGAGGAAGAAGAACCAAGTGAAAAATCAGAGGCCACCTACATGACCATGCA CCCAGTTTGGCCTTCTCTGAGGTCAGATCGGAACAACTCACTTGAAAAAAAGTCAGGTGGGG ${\tt GAATGCCAAAAACACAGCAAGCCTTT} {\color{red}{\textbf{TGA}}} {\color{red}{\textbf{GAAGAATGGAAGTCCCTTCATCTCAGCAGCGG}}$ TGGAGACTCTCTCCTGTGTGTGTCCTGGGCCACTCTACCAGTGATTTCAGACTCCCGCTCTC CCAGCTGTCCTCCTGTCTCATTGTTTGGTCAATACACTGAAGATGGAGAATTTGGAGCCTGG CAGAGAGACTGGACAGCTCTGGAGGAACAGGCCTGCTGAGGGGAGGGGAGCATGGACTTGGC CTCTGGAGTGGGACACTGGCCCTGGGAACCAGGCTGAGCTGAGTGGCCTCAAACCCCCCGTT GGATCAGACCCTCCTGTGGGCAGGGTTCTTAGTGGATGAGTTACTGGGAAGAATCAGAGATA AAAACCAACCCAAATCAA

FIGURE 304

MFCPLKLILLPVLLDYSLGLNDLNVSPPELTVHVGDSALMGCVFQSTEDKCIFKIDWTLSPG
EHAKDEYVLYYYSNLSVPIGRFQNRVHLMGDILCNDGSLLLQDVQEADQGTYICEIRLKGES
QVFKKAVVLHVLPEEPKELMVHVGGLIQMGCVFQSTEVKHVTKVEWIFSGRRAKEEIVFRYY
HKLRMSVEYSQSWGHFQNRVNLVGDIFRNDGSIMLQGVRESDGGNYTCSIHLGNLVFKKTIV
LHVSPEEPRTLVTPAALRPLVLGGNQLVIIVGIVCATILLLPVLILIVKKTCGNKSSVNSTV
LVKNTKKTNPEIKEKPCHFERCEGEKHIYSPIIVREVIEEEEPSEKSEATYMTMHPVWPSLR
SDRNNSLEKKSGGGMPKTQQAF

FIGURE 305

CTATGAAGAAGCTTCCTGGAAAACAATAAGCAAAGGAAAACAAATGTGTCCCATCTCACATG GTTCTACCCTACTAAAGACAGGAAGATCATAAACTGACAGATACTGAAATTGTAAGAGTTGG AAACTACATTTTGCAAAGTCATTGAACTCTGAGCTCAGTTGCAGTACTCGGGAAGCCATCCA GGATGAAGATGGATACATCACCTTAAATATTAAAACTCGGAAACCAGCTCTCGTCTCCGTTG GCCCTGCATCCTCCTGGTGGCGTGTGATGGCTTTGATTCTGCTGATCCTGTGCGTGGGG ATGGTTGTCGGGCTGGTGGCTCTGGGGATTTGGTCTGTCATGCAGCGCAATTACCTACAAGA TGAGAATGAAAATCGCACAGGAACTCTGCAACAATTAGCAAAGCGCTTCTGTCAATATGTGG TAAAACAATCAGAACTAAAGGGCACTTTCAAAGGTCATAAATGCAGCCCCTGTGACACAAAC TGGAGATATTATGGAGATAGCTGCTATGGGTTCTTCAGGCACAACTTAACATGGGAAGAGAG TAAGCAGTACTGCACTGACATGAATGCTACTCTCCTGAAGATTGACAACCGGAACATTGTGG AGTACATCAAAGCCAGGACTCATTTAATTCGTTGGGTCGGATTATCTCGCCAGAAGTCGAAT GAGGTCTGGAAGTGGGAGGATGGCTCGGTTATCTCAGAAAATATGTTTGAGTTTTTGGAAGA TGGAAAAGGAAATATGAATTGTGCTTATTTTCATAATGGGAAAATGCACCCTACCTTCTGTG AGAACAAACATTATTTAATGTGTGAGAGGAAGGCTGGCATGACCAAGGTGGACCAACTACCT **TAA**TGCAAAGAGGTGGACAGGATAACACAGATAAGGGCTTTATTGTACAATAAAAGATATGT ATGAATGCATCAGTAGCTGAAAAAAAAAAAAAAA

FIGURE 306

MQDEDGYITLNIKTRKPALVSVGPASSSWWRVMALILLILCVGMVVGLVALGIWSVMQRNYL QDENENRTGTLQQLAKRFCQYVVKQSELKGTFKGHKCSPCDTNWRYYGDSCYGFFRHNLTWE ESKQYCTDMNATLLKIDNRNIVEYIKARTHLIRWVGLSRQKSNEVWKWEDGSVISENMFEFL EDGKGNMNCAYFHNGKMHPTFCENKHYLMCERKAGMTKVDQLP

FIGURE 307

CCCACGCGTCCGCGCAGTCGCCAGTTCTGCCTCCGCCTGCCAGTCTCGCCCGCGATCCCGG CGCCGGAGGAGCTCGGACGCATGCTGAGCCCCCTCCTTTGCTGAAGCCCGAGTGCGGAGAA GCCCGGGCAAACGCAGGCTAAGGAGACCAAAGCGGCGAAGTCGCGAGACAGCGGACAAGCAG CGTCGTGGCC**ATG**GCGGCGGCTATCGCCAGCTCGCTCATCCGTCAGAAGAGGCCAAGCCCGCG AGCGCGAGAAATCCAACGCCTGCAAGTGTGTCAGCAGCCCCAGCAAAGGCAAGACCAGCTGC GACAAAAACAAGTTAAATGTCTTTTCCCGGGTCAAACTCTTCGGCTCCAAGAAGAGGCGCAG AAGAAGACCAGAGCCTCAGCTTAAGGGTATAGTTACCAAGCTATACAGCCGACAAGGCTACC ACTTGCAGCTGCAGGCGGATGGAACCATTGATGGCACCAAAGATGAGGACAGCACTTACACT CTGTTTAACCTCATCCCTGTGGGTCTGCGAGTGGTGGCTATCCAAGGAGTTCAAACCAAGCT GTACTTGGCAATGAACAGTGAGGGATACTTGTACACCTCGGAACTTTTCACACCTGAGTGCA AATTCAAAGAATCAGTGTTTGAAAATTATTATGTGACATATTCATCAATGATATACCGTCAG CAGCAGTCAGGCCGAGGGTGGTATCTGGGTCTGAACAAAGAAGGAGAGATCATGAAAGGCAA CCATGTGAAGAAGAACAAGCCTGCAGCTCATTTTCTGCCTAAACCACTGAAAGTGGCCATGT ACAAGGAGCCATCACTGCACGATCTCACGGAGTTCTCCCGATCTGGAAGCGGGACCCCAACC AAGAGCAGAAGTGTCTCTGGCGTGCTGAACGGAGGCAAATCCATGAGCCACAATGAATCAAC $G{f TAG}$ CCAGTGAGGGCAAAAGAAGGGCTCTGTAACAGAACCTTACCTCCAGGTGCTGTTGAAT CAGAGTTCACTATTCTATCTGCCATTAGACCTTCTTATCATCCATACTAAAGC

FIGURE 308

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA28498

><subunit 1 of 1, 245 aa, 1 stop

><MW: 27564, pI: 10.18, NX(S/T): 1

MAAAIASSLIRQKRQAREREKSNACKCVSSPSKGKTSCDKNKLNVFSRVKLFGSKKRRRRP EPQLKGIVTKLYSRQGYHLQLQADGTIDGTKDEDSTYTLFNLIPVGLRVVAIQGVQTKLYLA MNSEGYLYTSELFTPECKFKESVFENYYVTYSSMIYRQQQSGRGWYLGLNKEGEIMKGNHVK KNKPAAHFLPKPLKVAMYKEPSLHDLTEFSRSGSGTPTKSRSVSGVLNGGKSMSHNEST

N-glycosylation site.

amino acids 242-246

Glycosaminoglycan attachment site.

amino acids 165-169, 218-222

Tyrosine kinase phosphorylation site.

amino acids 93-100

N-myristoylation site.

amino acids 87-93, 231-237

ATP/GTP-binding site motif A (P-loop).

amino acids 231-239

HBGF/FGF family proteins

amino acids 78-94, 102-153

FIGURE 309

CCAGGATGGAGCTGGGGCCTGTATAGCCATATTATTGTTCTATGCTACTAGACATGGGGGGG ACTTGGTGAAAAAGGTATTATCCAGCCAGAGGGTCTGGGAGCCCTGTCTTACTGAACCTGGG CAACCTGGATATTCTGAGACATATTTTGGGGGGGATTTCAGTGAAAAAAGTGGGGGATCCCCT CCCCAGTAGGGGTGGGATGAGCGAATATTCCCAAAGCTAAAGTCCCACACCCTGTAGATTAC AAGAGTGGATTTGGCAGGAGTGTGCCCCAAAATACAGTGGAAAGGTGCCTGAAGATATTTAA GAGAGGAGGGAAAGGGGACGTTTTCAATAGGAGGCAAAACTCGAGGGTGGGATCCACTGAGG AGTACATAGGCTGCTGGATCTGGTGGAGCCAGCACTGGGCCCACGGGTGGTAACTGGCTGCT CGAGTCGGGGCCTGAGCGTCAAGAGCATGCCCTAGTGAGCGGGCTCCTCTGGGGGAGCCCAG CGCAGCGGCGCGTGTGTCCCCGCGGCACCAAGTCCCTTTGCCAGAAGCAGCTCCTCATCCTG CTGTCCAAGGTGCGACTGTGCGGGGGGGGGCCGGCCGGGCCGGACCGCGGCCCGGAGCCTCA GCTCAAAGGCATCGTCACCAAACTGTTCTGCCGCCAGGGTTTCTACCTCCAGGCGAATCCCG ACGGAAGCATCCAGGGCACCCCAGAGGATACCAGCTCCTTCACCCACTTCAACCTGATCCCT GTGGGCCTCCGTGTGGTCACCATCCAGAGCGCCAAGCTGGGTCACTACATGGCCATGAATGC $\tt TGAGGGACTGCTCTACAGTTCGCCGCATTTCACAGCTGAGTGTCGCTTTAAGGAGTGTCTCT$ TTGAGAATTACTACGTCCTGTACGCCTCTGCTCTACCGCCAGCGTCGTTCTGGCCGGGCC TGGTACCTCGGCCTGGACAAGGAGGGCCAGGTCATGAAGGGAAACCGAGTTAAGAAGACCAA GGCAGCTGCCCACTTTCTGCCCAAGCTCCTGGAGGTGGCCATGTACCAGGAGCCTTCTCTCC ACAGTGTCCCCGAGGCCTCCCCTTCCAGTCCCCCTGCCCCCTGAAATGTAGTCCCTGGACTG GAGGTTCCCTGCACTCCCAGTGAGCCAGCCACCACCACCACCTGT

FIGURE 310

MAALASSLIRQKREVREPGGSRPVSAQRRVCPRGTKSLCQKQLLILLSKVRLCGGRPARPDR GPEPQLKGIVTKLFCRQGFYLQANPDGSIQGTPEDTSSFTHFNLIPVGLRVVTIQSAKLGHY MAMNAEGLLYSSPHFTAECRFKECVFENYYVLYASALYRQRRSGRAWYLGLDKEGQVMKGNR VKKTKAAAHFLPKLLEVAMYQEPSLHSVPEASPSSPPAP

Tyrosine kinase phosphorylation site:

amino acids 199-207

N-myristoylation sites:

amino acids 54-60, 89-95, 131-137

HBGF/FGF family signature:

amino acids 131-155

FIGURE 311

FIGURE 312

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA28503

><subunit 1 of 1, 247 aa, 1 stop

><MW: 27702, pI: 10.36, NX(S/T): 2

MAAAIASGLIRQKRQAREQHWDRPSASRRRSSPSKNRGLCNGNLVDIFSKVRIFGLKKRRLR RQDPQLKGIVTRLYCRQGYYLQMHPDGALDGTKDDSTNSTLFNLIPVGLRVVAIQGVKTGLY IAMNGEGYLYPSELFTPECKFKESVFENYYVIYSSMLYRQQESGRAWFLGLNKEGQAMKGNR VKKTKPAAHFLPKPLEVAMYREPSLHDVGETVPKPGVTPSKSTSASAIMNGGKPVNKSKTT

N-glycosylation site.

amino acids 100-104, 242-246

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 28-32, 29-33

Tyrosine kinase phosphorylation site.

amino acids 199-207

N-myristoylation site.

amino acids 38-44, 89-95, 118-124, 122-128, 222-228

HBGF/FGF family proteins.

amino acids 104-155, 171-198

FIGURE 313

GAAGGATGCAGGACGCAGCTTTCTCCTGGAACCGAACGCAATGGATAAACTGATTGTGCAAGAGAAGAAGAAGA ACGAAGCTTTTTCTTGTGAGCCCTGGATCTTAACACAAATGTGTATATGTGCACACAGGGAGCATTCAAGAATG CTCCCCACCCCAAAAAAAAGGATGATTGGAAATGAAGAACCGAGGATTCACAAAGAAAAAAGTATGTTCATTT $\tt TTCTCTATAAAGGAGAAAGTGAGCCAAGGAGATATTTTTGGAATGAAAAGTTTGGGGGCTTTTTTAGTAAAGTAA$ AGAACTGGTGTGGTGGTGTTTTCCTTTCTTTTGAATTTCCCACAAGAGGAGAGGAAATTAATAATACATCTGC CAGTTGGATTTGTGCCTATGTTGACTAAAATTGACGGATAATTGCAGTTGGATTTTTCTTCATCAACCTCCTTT TTTTTAAATTTTTATTCCTTTTGGTATCAAGATCATGCGTTTTCTCTTGTTCTTAACCACCTGGATTTCCATCT GGATGTTGCTGTGATCAGTCTGAAATACAACTGTTTGAATTCCAGAAGGACCAACACCAGATAAATTATGA<u>ATG</u> TTGAACAAGATGACCTTACATCCACAGCAGATAATGATAGGTCCTAGGTTTAACAGGGCCCTATTTGACCCCCT GCTTGTGGTGCTGCTGGCTCTTCAACTTCTTGTGGTGGCTGGTCTGGTGCGGGCTCAGACCTGCCCTTCTGTGT GCTCCTGCAGCAACCAGTTCAGCAAGGTGATTTGTGTTCGGAAAAACCTGCGTGAGGTTCCGGATGGCATCTCC ACCAACACACGGCTGCTGAACCTCCATGAGAACCAAATCCAGATCATCAAAGTGAACAGCTTCAAGCACTTGAG GCACTTGGAAATCCTACAGTTGAGTAGGAACCATATCAGAACCATTGAAATTGGGGCTTTCAATGGTCTGGCGA ACCTCAACACTCTGGAACTCTTTGACAATCGTCTTACTACCATCCCGAATGGAGCTTTTGTATACTTGTCTAAA CTGAAGGAGCTCTGGTTGCGAAACAACCCCATTGAAAGCATCCCTTCTTATGCTTTTAACAGAATTCCTTCTTT GCGCCGACTAGACTTAGGGGAATTGAAAAGACTTTCATACATCTCAGAAGGTGCCTTTGAAGGTCTGTCCAACT TGAGGTATTTGAACCTTGCCATGTGCAACCTTCGGGAAATCCCTAACCTCACACCGCTCATAAAACTAGATGAG CTGGATCTTTCTGGGAATCATTTATCTGCCATCAGGCCTGGCTCTTTCCAGGGTTTGATGCACCTTCAAAAACT GTGGATGATACAGTCCCAGATTCAAGTGATTGAACGGAATGCCTTTGACAACCTTCAGTCACTAGTGGAGATCA ACCTGGCACACAATAATCTAACATTACTGCCTCATGACCTCTTCACTCCCTTGCATCATCTAGAGCGGATACAT TTACATCACAACCCTTGGAACTGTAACTGTGACATACTGTGGCTCAGCTGGTGGATAAAAGACATGGCCCCCTC GAACACAGCTTGTTGTGCCCGGTGTAACACTCCTCCCAATCTAAAGGGGAGGTACATTGGAGAGCTCGACCAGA ATTACTTCACATGCTATGCTCCGGTGATTGTGGAGCCCCCTGCAGACCTCAATGTCACTGAAGGCATGGCAGCT GAGCTGAAATGTCGGGCCTCCACATCCCTGACATCTGTATCTTGGATTACTCCAAATGGAACAGTCATGACACA TGGGGCGTACAAAGTGCGGATAGCTGTGCTCAGTGATGGTACGTTAAATTTCACAAATGTAACTGTGCAAGATA CAGGCATGTACACATGTATGGTGAGTAATTCCGTTGGGAATACTACTGCTTCAGCCACCCTGAATGTTACTGCA GCAACCACTACTCCTTTCTCTTACTTTTCAACCGTCACAGTAGAGACTATGGAACCGTCTCAGGATGAGGCACG GACCACAGATAACAATGTGGGTCCCACTCCAGTGGTCGACTGGGAGACCACCAATGTGACCACCTCTCTCACAC CACAGAGCACAAGGTCGACAGAGAAAACCTTCACCATCCCAGTGACTGATATAAACAGTGGGATCCCAGGAATT GATGAGGTCATGAAGACTACCAAAATCATCATTGGGTGTTTTGTGGCCATCACACTCATGGCTGCAGTGATGCT GGTCATTTTCTACAAGATGAGGAAGCAGCACCATCGGCAAAACCATCACGCCCCAACAAGGACTGTTGAAATTA TTAATGTGGATGATGAGATTACGGGAGACACACCCATGGAAAGCCACCTGCCCATGCCTATCGAGCATGAG TGATCTAAAGCAGACAAAAA

FIGURE 314

MLNKMTLHPQQIMIGPRFNRALFDPLLVVLLALQLLVVAGLVRAQTCPSVCSCSNQFSKVIC
VRKNLREVPDGISTNTRLLNLHENQIQIIKVNSFKHLRHLEILQLSRNHIRTIEIGAFNGLA
NLNTLELFDNRLTTIPNGAFVYLSKLKELWLRNNPIESIPSYAFNRIPSLRRLDLGELKRLS
YISEGAFEGLSNLRYLNLAMCNLREIPNLTPLIKLDELDLSGNHLSAIRPGSFQGLMHLQKL
WMIQSQIQVIERNAFDNLQSLVEINLAHNNLTLLPHDLFTPLHHLERIHLHHNPWNCNCDIL
WLSWWIKDMAPSNTACCARCNTPPNLKGRYIGELDQNYFTCYAPVIVEPPADLNVTEGMAAE
LKCRASTSLTSVSWITPNGTVMTHGAYKVRIAVLSDGTLNFTNVTVQDTGMYTCMVSNSVGN
TTASATLNVTAATTTPFSYFSTVTVETMEPSQDEARTTDNNVGPTPVVDWETTNVTTSLTPQ
STRSTEKTFTIPVTDINSGIPGIDEVMKTTKIIIGCFVAITLMAAVMLVIFYKMRKQHHRQN
HHAPTRTVEIINVDDEITGDTPMESHLPMPAIEHEHLNHYNSYKSPFNHTTTVNTINSIHSS
VHEPLLIRMNSKDNVQETQI

Signal sequence:

amino acids 1-44

Transmembrane domain:

amino acids 523-543

N-qlycosylation site.

amino acids 278-282, 364-368, 390-394, 412-416, 415-419, 434-438, 442-446, 488-492, 606-610

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 183-187

Casein kinase II phosphorylation site.

amino acids 268-272, 417-421, 465-469, 579-583, 620-624

N-myristoylation site.

amino acids 40-46, 73-79, 118-124, 191-197, 228-234, 237-243, 391-397, 422-428, 433-439, 531-537

FIGURE 315

GCGCCGGGAGCCCATCTGCCCCCAGGGGCCACGGGGCGGGGGCCGGCTCCCGGCCCGGCACAT GGCTGCAGCCACCTCGCGCGCACCCCGAGGCGCCCCAGCTCGCCCGAGGTCCGTCGGA GGCGCCCGGCCCCCGGAGCCAAGCAGCAACTGAGCGGGGAAGCGCCCGCGTCCGGGGATC $\mathsf{GGG} \underline{\mathbf{ATG}} \mathsf{TCCCTCCTTCTCTCTTGCTAGTTTCCTACTATGTTGGAACCTTGGGGACTCA$ CACTGAGATCAAGAGAGTGGCAGAGGAAAAGGTCACTTTGCCCTGCCACCATCAACTGGGGC TTCCAGAAAAAGACACTCTGGATATTGAATGGCTGCTCACCGATAATGAAGGGAACCAAAAA GTGGTGATCACTTACTCCAGTCGTCATGTCTACAATAACTTGACTGAGGAACAGAAGGGCCG AGTGGCCTTTGCTTCCAATTTCCTGGCAGGAGATGCCTCCTTGCAGATTGAACCTCTGAAGC CCAGTGATGAGGGCCGGTACACCTGTAAGGTTAAGAATTCAGGGCGCTACGTGTGGAGCCAT GTCATCTTAAAAGTCTTAGTGAGACCATCCAAGCCCAAGTGTGAGTTGGAAGGAGAGCTGAC AGAAGGAAGTGACCTGACTTTGCAGTGTGAGTCATCCTCTGGCACAGAGCCCATTGTGTATT ACTGGCAGCGAATCCGAGAGAAAGAGGGGAGGATGAACGTCTGCCTCCCAAATCTAGGATT GACTACAACCACCCTGGACGAGTTCTGCTGCAGAATCTTACCATGTCCTACTCTGGACTGTA CCAGTGCACAGCAGCAACGAAGCTGGGAAGGAAAGCTGTGTGGGGGGAGTAACTGTACAGT ATGTACAAAGCATCGGCATGGTTGCAGGAGCAGTGACAGGCATAGTGGCTGGAGCCCTGCTG GAGACCTAATGAAATTCGAGAAGATGCTGAAGCTCCAAAAGCCCGTCTTGTGAAACCCAGCT CCTCTTCCTCAGGCTCTCGGAGCTCACGCTCTGGTTCTTCCTCCACTCGCTCCACAGCAAAT ACGGTC TCAATGGACTTGACTCCCACGCTTTCCTAGGAGTCAGGGTCTTTGGACTC TTCTCGTCATTGGAGCTCAAGTCACCAGCCACACAACCAGATGAGAGGTCATCTAAGTAGCA GTGAGCATTGCACGGAACAGATTCAGATGAGCATTTTCCTTATACAATACCAAACAAGCAAA AGGATGTAAGCTGATTCATCTGTAAAAAGGCATCTTATTGTGCCTTTAGACCAGAGTAAGGG AAAGCAGGAGTCCAAATCTATTTGTTGACCAGGACCTGTGGTGAGAAGGTTGGGGAAAGGTG AGGTGAATATACCTAAAACTTTTAATGTGGGATATTTTGTATCAGTGCTTTGATTCACAATT TTCAAGAGGAAATGGGATGCTGTTTGTAAATTTTCTATGCATTTCTGCAAACTTATTGGATT ATTAGTTATTCAGACAGTCAAGCAGAACCCACAGCCTTATTACACCTGTCTACACCATGTAC TGAGCTAACCACTTCTAAGAAACTCCAAAAAAGGAAACATGTGTCTTCTATTCTGACTTAAC TTCATTTGTCATAAGGTTTGGATATTAATTTCAAGGGGAGTTGAAATAGTGGGAGATGGAGA AGAGTGAATGAGTTTCTCCCACTCTATACTAATCTCACTATTTGTATTGAGCCCAAAATAAC TATGAAAGGAGACAAAAATTTGTGACAAAGGATTGTGAAGAGCTTTCCATCTTCATGATGTT ATGAGGATTGTTGACAAACATTAGAAATATATAATGGAGCAATTGTGGATTTCCCCTCAAAT CAGATGCCTCTAAGGACTTTCCTGCTAGATATTTCTGGAAGGAGAAAATACAACATGTCATT TATCAACGTCCTTAGAAAGAATTCTTCTAGAGAAAAAGGGATCTAGGAATGCTGAAAGATTA CCCAACATACCATTATAGTCTCTTCTTTCTGAGAAAATGTGAAACCAGAATTGCAAGACTGG TGGTGCCAGGCACCTGTAGGAAAATCCAGCAGGTGGAGGTTGCAGTGAGCCGAGATTATGCC ATTGCACTCCAGCCTGGGTGACAGAGCGGGACTCCGTCTC

FIGURE 316

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45419

><subunit 1 of 1, 373 aa, 1 stop

><MW: 41281, pI: 8.33, NX(S/T): 3

MSLLLLLLVSYYVGTLGTHTEIKRVAEEKVTLPCHHQLGLPEKDTLDIEWLLTDNEGNQKV VITYSSRHVYNNLTEEQKGRVAFASNFLAGDASLQIEPLKPSDEGRYTCKVKNSGRYVWSHV ILKVLVRPSKPKCELEGELTEGSDLTLQCESSSGTEPIVYYWQRIREKEGEDERLPPKSRID YNHPGRVLLQNLTMSYSGLYQCTAGNEAGKESCVVRVTVQYVQSIGMVAGAVTGIVAGALLI FLLVWLLIRRKDKERYEEEERPNEIREDAEAPKARLVKPSSSSSGSRSSRSGSSSTRSTANS ASRSQRTLSTDAAPQPGLATQAYSLVGPEVRGSEPKKVHHANLTKAETTPSMIPSQSRAFQTV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 232-251

FIGURE 317

 ${\tt CC}{ ext{ATG}}{\tt GCGCTCCTGCTGCTGCTGCGGGGGGTAGTGGGATTTCGCCAGAAGTTTGAGTATCACTACT$ CCTGAAGAGATGATTGAAAAAGCCAAAGGGGAAACTGCCTATCTGCCATGCAAATTTACGCTTAGTCCCGAAGA CCAGGGACCGCTGGACATCGAGTGGCTGATATCACCAGCTGATAATCAGAAGGTGGATCAAGTGATTATTTTAT ATTCTGGAGACAAAATTTATGATGACTACTATCCAGATCTGAAAGGCCGAGTACATTTTACGAGTAATGATCTC AAATCTGGTGATGCATCAATAAATGTAACGAATTTACAACTGTCAGATATTGGCACATATCAGTGCAAAGTGAA AAAAGCTCCTGGTGTTGCAAATAAGAAGATTCATCTGGTAGTTCTTGTTAAGCCTTCAGGTGCGAGATGTTACG TTGATGGATCTGAAGAAATTGGAAGTGACTTTAAGATAAAATGTGAACCAAAAGAAGGTTCACTTCCATTACAG TATGAGTGGCAAAAATTGTCTGACTCACAGAAAATGCCCCACTTCATGGTTAGCAGAAATGACTTCATCTGTTAT ATCTGTAAAAAATGCCTCTTCTGAGTACTCTGGGACATACAGCTGTACAGTCAGAAACAGAGTGGGCTCTGATC AGTGCCTGTTGCGTCTAAACGTTGTCCCTCCTTCAAATAAAGCTGGACTAATTGCAGGAGCCATTATAGGAACT GGAAGTTCATCACGATATCAGGGAAGATGTGCCACCTCCAAAGAGCCGTACGTCCACTGCCAGAAGCTACATCG GCAGTAATCATTCATCCCTGGGGTCCATGTCTCCTTCCAACATGGAAGGATATTCCAAGACTCAGTATAACCAA GTACCAAGTGAAGACTTTGAACGCACTCCTCAGAGTCCGACTCTCCCACCTGCTAAGTTCAAGTACCCTTACAA ${ t GACTGATGGAATTACAGTTGTA}$ AGGCCTCTAGTAAAGACTTAAATGTTTTTTAAAAAAAGCACAAGGCACAGAGATTAGAGCAGCTGTAAGAACAC ATCTACTTTATGCAATGGCATTAGACATGTAAGTCAGATGTCATGTCAAAATTAGTACGAGCCAAATTCTTTGT TAAAAAACCCTATGTATAGTGACACTGATAGTTAAAAGATGTTTTATTATATTTTCAATAACTACCACTAACAA ATTTTTAACTTTTCATATGCATATTCTGATATGTGGTCTTTTAGGAAAAGTATGGTTAATAGTTGATTTTTCAA AGGAAATTTTAAAATTCTTACGTTCTGTTTAATGTTTTTGCTATTTAGTTAAATACATTGAAGGGAAATACCCG TAAAGTAAATTCTCAAAGGTGCTAGAACAAATCGTCCACTTCTACAGTGTTCTCGTATCCAACAGAGTTGATGC ACAATATATAAATACTCAAGTCCAATATTAAAAACTTAGGCACTTGACTAACTTTAATAAAATTTCTCAAACTA TATCAATATCTAAAGTGCATATATTTTTTAAGAAAGATTATTCTCAATAACTTCTATAAAAATAAGTTTGATGG ${\tt TTTGGCCCATCTAACTTCACTACTATTAGTAAGAACTTTTAACTTTTAATGTGTAGTAAGGTTTATTCTACCTT}$ TTTCTCAACATGACACCAACACAATCAAAAACGAAGTTAGTGAGGTGCTAACATGTGAGGATTAATCCAGTGAT ${\tt TCCGGTCACAATGCATTCCAGGAGGAGGTACCCATGTCACTGGAATTGGGCGATATGGTTTATTTTTTCTTCCCC}$ TGATTTGGATAACCAAATGGAACAGGAGGAGGATAGTGATTCTGATGGCCATTCCCTCGATACATTCCTGGCTT TTTTCTGGGCAAAGGGTGCCACATTGGAAGAGGTGGAAATATAAGTTCTGAAATCTGTAGGGAAGAGAACACAT TAAGTTAATTCAAAGGAAAAAATCATCATCTATGTTCCAGATTTCTCATTAAAGACAAAGTTACCCACAACACT GAGATCACATCTAAGTGACACTCCTATTGTCAGGTCTAAATACATTAAAAACCTCATGTGTAATAGGCGTATAA TGTATAACAGGTGACCAATGTTTTCTGAATGCATAAAGAAATGAATAAACTCAAACACAGTACTTCCTAAACAA CTTCAACCAAAAAAGACCAAAACATGGAACGAATGGAAGCTTGTAAGGACATGCTTGTTTTAGTCCAGTGGTTT CCACAGCTGGCTAAGCCAGGAGTCACTTGGAGGCTTTTAAATACAAAACATTGGAGCTGGAGGCCATTATCCTT AGCAAACTAATGCAGAAACAGAAAATCAACTACCGCATGTTCTCACTTATAAGTGGGAGGTAATGATAAGAACT TATGAACACAAAGAAGGAAACAATAGACATTGGAGTCTATTTGAGAGGGGAGGGTGGGAGAAGGAAAAGGAGCA GAAAAGATAACTATTGAGTACTGCCTTCACACCTGGGTGATGAAATAATATGTACAACAAATCCCTGTGACACA AAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 318

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA82361

><subunit 1 of 1, 352 aa, 1 stop

><MW: 38938, pI: 7.86, NX(S/T): 3

MALLLCFVLLCGVVDFARSLSITTPEEMIEKAKGETAYLPCKFTLSPEDQGPLDIEWLISPA
DNQKVDQVIILYSGDKIYDDYYPDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQCKVKK
APGVANKKIHLVVLVKPSGARCYVDGSEEIGSDFKIKCEPKEGSLPLQYEWQKLSDSQKMPT
SWLAEMTSSVISVKNASSEYSGTYSCTVRNRVGSDQCLLRLNVVPPSNKAGLIAGAIIGTLL
ALALIGLIIFCCRKKRREEKYEKEVHHDIREDVPPPKSRTSTARSYIGSNHSSLGSMSPSNM
EGYSKTQYNQVPSEDFERTPQSPTLPPAKFKYPYKTDGITVV

Signal sequence.

amino acids 1-19

Transmembrane domain:

amino acids 236-257

N-glycosylation sites.

amino acids 106-110, 201-205, 298-302

Tyrosine kinase phosphorylation sites.

amino acids 31-39, 78-85, 262-270

N-myristoylation sites.

amino acids 116-122, 208-214, 219-225, 237-243, 241-247, 245-251, 296-302

Myelin P0 protein.

amino acids 96-125

FIGURE 319

CTCAAGCATCACTTACAGGACCAGAGGGACAAGACATGACTGTGATGAGGAGCTGCTTTCGC CAATTTAACACCAAGAAGAATTGAGGCTGCTTGGGAGGAAGGCCAGGAGGAACACGAGACTG AGAG**ATG**AATTTTCAACAGAGGCTGCAAAGCCTGTGGACTTTAGCCAGACCCTTCTGCCCTC CTTTGCTGGCGACAGCCTCTCAAATGCAGATGGTTGTGCTCCCTTGCCTGGGTTTTACCCTG CTTCTCTGGAGCCAGGTATCAGGGGCCCAGGGCCAAGAATTCCACTTTGGGCCCTGCCAAGT GAAGGGGGTTGTTCCCCAGAAACTGTGGGAAGCCTTCTGGGCTGTGAAAGACACTATGCAAG CTCAGGATAACATCACGAGTGCCCGGCTGCTGCAGCAGGAGGTTCTGCAGAACGTCTCGGAT GCTGAGAGCTGTTACCTTGTCCACACCCTGCTGGAGTTCTACTTGAAAACTGTTTTCAAAAA CCACCACAATAGAACAGTTGAAGTCAGGACTCTGAAGTCATTCTCTACTCTGGCCAACAACT TTGTTCTCATCGTGTCACAACTGCAACCCAGTCAAGAAAATGAGATGTTTTCCATCAGAGAC AGTGCACACAGGCGGTTTCTGCTATTCCGGAGAGCATTCAAACAGTTGGACGTAGAAGCAGC TCTGACCAAAGCCCTTGGGGAAGTGGACATTCTTCTGACCTGGATGCAGAAATTCTACAAGC TC**TGA**ATGTCTAGACCAGGACCTCCCTCCCCTGGCACTGGTTTGTTCCCTGTGTCATTTCA AACAGTCTCCCTTCCTATGCTGTTCACTGGACACTTCACGCCCTTGGCCATGGGTCCCATTC TTGGCCCAGGATTATTGTCAAAGAAGTCATTCTTTAAGCAGCGCCAGTGACAGTCAGGGAAG AATTAATGTCAGTATTTCAACTGAAGTTCTATTTATTTGTGAGACTGTAAGTTACATGAAGG CAGCAGAATATTGTGCCCCATGCTTCTTTACCCCTCACAATCCTTGCCACAGTGTGGGGCAG GTTAAAAAACAGAGAGGGATGCTTGGATGTAAAACTGAACTTCAGAGCATGAAAATCACACT TAAACGATAAAATGTGGATTAAAGTGCCCAGCACAAAGCAGATCCTCAATAAACATTTCATT TATCCTAGTCATTCTTCCCTAATCTTCCACTTGAGTGTCAAGCTGACCTTGCTGATGGTGAC ATTGCACCTGGATGTACTATCCAATCTGTGATGACATTCCCTGCTAATAAAAGACAACATAA CTCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 320

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA88002

><subunit 1 of 1, 206 aa, 1 stop

><MW: 23799, pI: 9.12, NX(S/T): 3

MNFQQRLQSLWTLARPFCPPLLATASQMQMVVLPCLGFTLLLWSQVSGAQGQEFHFGPCQVK GVVPQKLWEAFWAVKDTMQAQDNITSARLLQQEVLQNVSDAESCYLVHTLLEFYLKTVFKNH HNRTVEVRTLKSFSTLANNFVLIVSQLQPSQENEMFSIRDSAHRRFLLFRRAFKQLDVEAAL TKALGEVDILLTWMQKFYKL

Signal sequence:

amino acids 1-42

N-glycosylation sites.

amino acids 85-89, 99-103, 126-130

FIGURE 321

FIGURE 322

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA92282

><subunit 1 of 1, 177 aa, 1 stop

><MW: 20452, pI: 8.00, NX(S/T): 2

MKLQCVSLWLLGTILILCSVDNHGLRRCLISTDMHHIEESFQEIKRAIQAKDTFPNVTILST LETLQIIKPLDVCCVTKNLLAFYVDRVFKDHQEPNPKILRKISSIANSFLYMQKTLRQCQEQ RQCHCRQEATNATRVIHDNYDQLEVHAAAIKSLGELDVFLAWINKNHEVMFSA

Signal sequence:

amino acids 1-18

N-glycosylation sites.

amino acids 56-60, 135-139

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 102-106

N-myristoylation site.

amino acids 24-30

Actinin-type actin-binding domain signature 1.

amino acids 159-169

FIGURE 323

CCCGTGCCAAGAGTGACGTAAGTACCGCCTATAGAGTCTATAGGCCCACTTGGCTTCGTTAG AACGCGGCTACAATTAATACATAACCTTATGTATCATACACATACGATTTAGGTGACACTAT AGAATAACATCCACTTTGCCTTTCTCTCCACAGGTGTCCACTCCCAGGTCCAACTGCACCTC ${\tt GGTTCTATCGATAATCTCAGCACCAGCCACTCAGAGCAGGGCACG} {\color{red} {\bf ATG}} {\tt TTGGGGGCCCGCCT}$ CAGGCTCTGGGTCTGTGCCTTGTGCAGCGTCTGCAGCATGAGCGTCCTCAGAGCCTATCCCA ATGCCTCCCCACTGCTCGGCTCCAGCTGGGGTGGCCTGATCCACCTGTACACAGCCACAGCC AGGAACAGCTACCACCTGCAGATCCACAAGAATGGCCATGTGGATGGCGCACCCCATCAGAC CATCTACAGTGCCCTGATGATCAGATCAGAGGATGCTGGCTTTTGTGGTGATTACAGGTGTGA TGAGCAGAAGATACCTCTGCATGGATTTCAGAGGCAACATTTTTGGATCACACTATTTCGAC CCGGAGAACTGCAGGTTCCAACACCAGACGCTGGAAAACGGGTACGACGTCTACCACTCTCC TCAGTATCACTTCCTGGTCAGTCTGGGCCGGGCGAAGAGAGCCTTCCTGCCAGGCATGAACC CACCCCGTACTCCCAGTTCCTGTCCCGGAGGAACGAGATCCCCCTAATTCACTTCAACACC CCCATACCACGGCGCACACCCGGAGCGCCGAGGACGACTCGGAGCGGGACCCCCTGAACGT GCTGAAGCCCCGGGCCCGGATGACCCCGGCCCCGGCCTCCTGTTCACAGGAGCTCCCGAGCG ACGCACGCTGGGGGAACGGCCCGGAAGGCTGCCGCCCCTTCGCCAAGTTCATC<u>TAG</u>GGTCG CTGG

FIGURE 324

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA142238

><subunit 1 of 1, 251 aa, 1 stop

><MW: 27954, pI: 9.22, NX(S/T): 1

MLGARLRLWVCALCSVCSMSVLRAYPNASPLLGSSWGGLIHLYTATARNSYHLQIHKNGHVD GAPHQTIYSALMIRSEDAGFVVITGVMSRRYLCMDFRGNIFGSHYFDPENCRFQHQTLENGY DVYHSPQYHFLVSLGRAKRAFLPGMNPPPYSQFLSRRNEIPLIHFNTPIPRRHTRSAEDDSE RDPLNVLKPRARMTPAPASCSQELPSAEDNSPMASDPLGVVRGGRVNTHAGGTGPEGCRPFA KFI

Important features of the protein:

Signal peptide:

amino acids 1-24

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 175-179

N-myristoylation site.

amino acids 33-39, 100-106, 225-231, 229-235

HBGF/FGF family proteins

amino acids 73-124

FIGURE 325

GGAAAAGGTACCCGCGAGAGACAGCCAGCAGTTCTGTGGAGCAGCGGTGGCCGGCTAGG**ATG** GGCTGTCTCTGGGGTCTGGCCCCTTTTCTTCTTCTGCTGGGAGGTTGGGGTCTCTGG GAGCTCTGCAGGCCCCAGCACCCGCAGAGCAGACACTGCGATGACAACGGACGACACAGAAG TGCCCGCTATGACTCTAGCACCGGGCCACGCCGCTCTGGAAACTCAAACGCTGAGCGCTGAG ACCTCTTCTAGGGCCTCAACCCCAGCCGGCCCCATTCCAGAAGCAGAGACCAGGGGAGCCAA GAGAATTTCCCCTGCAAGAGAGACCAGGAGTTTCACAAAAACATCTCCCAACTTCATGGTGC TGATCGCCACCTCCGTGGAGACATCAGCCGCCAGTGGCAGCCCCGAGGGAGCTGGAATGACC ACAGTTCAGACCATCACAGGCAGTGATCCCGAGGAAGCCATCTTTGACACCCTTTGCACCGA TGACAGCTCTGAAGAGGCAAAGACACTCACAATGGACATATTGACATTGGCTCACACCTCCA CAGAAGCTAAGGGCCTGTCCTCAGAGAGCAGTGCCTCTTCCGACGGCCCCCATCCAGTCATC ACCCCGTCACGGGCCTCAGAGAGCAGCGCCTCTTCCGACGGCCCCCATCCAGTCATCACCCC GTCACGGGCCTCAGAGAGCAGCGCCTCTTCCGACGGCCCCCATCCAGTCATCACCCCGTCAT GGTCCCCGGGATCTGATGTCACTCTCCTCGCTGAAGCCCTGGTGACTGTCACAAACATCGAG GTTATTAATTGCAGCATCACAGAAATAGAAACAACAACTTCCAGCATCCCTGGGGCCTCAGA CATAGATCTCATCCCCACGGAAGGGGTGAAGGCCTCGTCCACCTCCGATCCACCAGCTCTGC CTGACTCCACTGAAGCAAAACCACACATCACTGAGGTCACAGCCTCTGCCGAGACCCTGTCC ACAGCCGGCACCACAGAGTCAGCTGCACCTCATGCCACGGTTGGGACCCCACTCCCCACTAA CAGCGCCACAGAAAGAGAAGTGACAGCACCCGGGGCCACGACCCTCAGTGGAGCTCTGGTCA CAGTTAGCAGGAATCCCCTGGAAGAACCTCAGCCCTCTCTGTTGAGACACCAAGTTACGTC AAAGTCTCAGGAGCAGCTCCGGTCTCCATAGAGGCTGGGTCAGCAGTGGGCAAAACAACTTC CTTTGCTGGGAGCTCTGCTTCCTCCTACAGCCCCTCGGAAGCCGCCCTCAAGAACTTCACCC CTTCAGAGACACCGACCATGGACATCGCAACCAAGGGGCCCTTCCCCACCAGCAGGGACCCT CTTCCTTCTGTCCCTCCGACTACAACCAACAGCAGCCGAGGGACGAACAGCACCTTAGCCAA GATCACAACCTCAGCGAAGACCACGATGAAGCCCCAACAGCCACGCCCACGACTGCCCGGAC ${\sf GAGGCCGACCACAGACG}$ ${\sf TGA}$ ${\sf GTGCAGGTGAAAATGGAGGTTTCCTCCTCCTGCGGCTGAGTG}$ TGGCTTCCCCGGAAGACCTCACTGACCCCAGAGTGGCAGAAAGGCTGATGCAGCAGCTCCAC CGGGAACTCCACGCCCACGCCCTCACTTCCAGGTCTCCTTACTGCGTGTCAGGAGAGGCTA ACGGACATCAGCTGCAGCCAGGCATGTCCCGTATGCCAAAAGAGGGTGCTGCCCCTAGCCTG GGCCCCCACCGACAGACTGCAGCTGCGTTACTGTGCTGAGAGGTACCCAGAAGGTTCCCATG AAGGGCAGCATGTCCAAGCCCCTAACCCCAGATGTGGCAACAGGACCCTCGCTCACATCCAC CGGAGTGTATGTATGGGGAGGGGCTTCACCTGTTCCCAGAGGTGTCCTTGGACTCACCTTGG CACATGTTCTGTGTTTCAGTAAAGAGAGACCTGATCACCCATCTGTGTGCTTCCATCCTGCA TTAAAATTCACTCAGTGTGGCCCAAAAAAA

FIGURE 326

MGCLWGLALPLFFFCWEVGVSGSSAGPSTRRADTAMTTDDTEVPAMTLAPGHAALETQTLSA
ETSSRASTPAGPIPEAETRGAKRISPARETRSFTKTSPNFMVLIATSVETSAASGSPEGAGM
TTVQTITGSDPEEAIFDTLCTDDSSEEAKTLTMDILTLAHTSTEAKGLSSESSASSDGPHPV
ITPSRASESSASSDGPHPVITPSRASESSASSDGPHPVITPSWSPGSDVTLLAEALVTVTNI
EVINCSITEIETTTSSIPGASDIDLIPTEGVKASSTSDPPALPDSTEAKPHITEVTASAETL
STAGTTESAAPHATVGTPLPTNSATEREVTAPGATTLSGALVTVSRNPLEETSALSVETPSY
VKVSGAAPVSIEAGSAVGKTTSFAGSSASSYSPSEAALKNFTPSETPTMDIATKGPFPTSRD
PLPSVPPTTTNSSRGTNSTLAKITTSAKTTMKPQQPRPRLPGRGRPQT

N-glycosylation sites:

amino acids 252-256, 445-449, 451-455

cAMP-and cGMP-dependent protein kinase phosphorylation site. amino acids 84-90

Casein kinase II phosphorylation sites.

amino acids 37-41, 108-112, 131-135, 133-137, 148-152, 165-169, 246-250, 254-258, 256-260, 269-273, 283-287, 333-337, 335-339, 404-408, 414-418, 431-435

N-myristoylation sites.

amino acids 2-8, 19-25, 117-123, 121-127, 232-238, 278-284, 314-320, 349-355, 386-392, 397-403, 449-455

ATP/GTP-binding site motif A (P-loop).

amino acids 385-393

FIGURE 327

GCGGAGCATCCGCTGCGGTCCTCGCCGAGACCCCCGCGCGGATTCGCCGGTCCTTCCCGCGG GCGCGACAGAGCTGTCCTCGCACCTGGATGGCAGCAGGGGCGCCGGGGTCCTCTCGACGCCA CCTTGACCTTTGAAGACCAAAACTAAACTGAAATTTAAA**ATG**TTCTTCGGGGGAGAAGGGAG CTTGACTTACACTTTGGTAATAATTTGCTTCCTGACACTAAGGCTGTCTGCTAGTCAGAATT GCCTCAAAAAGAGTCTAGAAGATGTTGTCATTGACATCCAGTCATCTCTTTCTAAGGGAATC AGAGGCAATGAGCCCGTATATACTTCAACTCAAGAAGACTGCATTAATTCTTGCTGTTCAAC AAAAAACATATCAGGGGACAAAGCATGTAACTTGATGATCTTCGACACTCGAAAAACAGCTA GACAACCCAACTGCTACCTATTTTTCTGTCCCAACGAGGAAGCCTGTCCATTGAAACCAGCA AAAGGACTTATGAGTTACAGGATAATTACAGATTTTCCATCTTTGACCAGAAATTTGCCAAG CCAAGAGTTACCCCAGGAAGATTCTCTCTTACATGGCCAATTTTCACAAGCAGTCACTCCCC TAGCCCATCATCACACAGATTATTCAAAGCCCACCGATATCTCATGGAGAGACACACTTTCT GCTCCTTGCTTATAAGGAAAAAGGCCATTCTCAGAGTTCACAATTTTCCTCTGATCAAGAAA TAGCTCATCTGCTGCCTGAAAATGTGAGTGCGCTCCCAGCTACGGTGGCAGTTGCTTCTCCA CATACCACCTCGGCTACTCCAAAGCCCGCCACCCTTCTACCCACCAATGCTTCAGTGACACC TTCTGGGACTTCCCAGCCACAGCTGGCCACCACAGCTCCACCTGTAACCACTGTCACTTCTC AGCCTCCACGACCCTCATTTCTACAGTTTTTTACACGGGCTGCGGCTACACTCCAAGCAATG GCTACAACAGCAGTTCTGACTACCACCTTTCAGGCACCTACGGACTCGAAAGGCAGCTTAGA AACCATACCGTTTACAGAAATCTCCAACTTAACTTTGAACACAGGGAATGTGTATAACCCTA CTGCACTTTCTATGTCAAATGTGGAGTCTTCCACTATGAATAAAACTGCTTCCTGGGAAGGT AGGGAGGCCAGTCCAGGCAGTTCCTCCCAGGGCAGTGTTCCAGAAAATCAGTACGGCCTTCC ATTTGAAAAATGGCTTCTTATCGGGTCCCTGCTCTTTGGTGTCCTGTTCCTGGTGATAGGCC TCGTCCTCCTGGGTAGAATCCTTTCGGAATCACTCCGCAGGAAACGTTACTCAAGACTGGAT TATTTGATCAATGGGATCTATGTGGACATC**TAA**GGATGGAACTCGGTGTCTCTTAATTCATT TAGTAACCAGAAGCCCAAATGCAATGAGTTTCTGCTGACTTGCTAGTCTTAGCAGGAGGTTG GCTCTGTTGCCCAGGCTGGAGTGCAGTAGCACGATCTCGGCTCTCACCGCAACCTCCGTCTC CTGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCTAAGTATCTGGGATTACAGGCATGTGCCA CCACACCTGGGTGATTTTTGTATTTTTAGTAGAGACGGGGTTTCACCATGTTGGTCAGGCTG GTCTCAAACTCCTGACCTAGTGATCCACCCTCCTCGGCCTCCCAAAGTGCTGGGATTACAGG CATGAGCCACCACAGCTGGCCCCCTTCTGTTTTATGTTTTGGTTTTTGAGAAGGAATGAAGTG GGAACCAAATTAGGTAATTTTGGGTAATCTGTCTCTAAAATATTAGCTAAAAACAAAGCTCT ATGTAAAGTAATAAAGTATAATTGCCATATAAATTTCAAAATTCAACTGGCTTTTATGCAAA GAAACAGGTTAGGACATCTAGGTTCCAATTCATTCACATTCTTGGTTCCAGATAAAATCAAC TGTTTATATCAATTTCTAATGGATTTGCTTTTCTTTTTATATGGATTCCTTTAAAACTTATT CCAGATGTAGTTCCTTCCAATTAAATATTTGAATAAATCTTTTGTTACTCAA

FIGURE 328

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45410

><subunit 1 of 1, 431 aa, 1 stop

><MW: 46810, pI: 6.45, NX(S/T): 6

MFFGGEGSLTYTLVIICFLTLRLSASQNCLKKSLEDVVIDIQSSLSKGIRGNEPVYTSTQED CINSCCSTKNISGDKACNLMIFDTRKTARQPNCYLFFCPNEEACPLKPAKGLMSYRIITDFP SLTRNLPSQELPQEDSLLHGQFSQAVTPLAHHHTDYSKPTDISWRDTLSQKFGSSDHLEKLF KMDEASAQLLAYKEKGHSQSSQFSSDQEIAHLLPENVSALPATVAVASPHTTSATPKPATLL PTNASVTPSGTSQPQLATTAPPVTTVTSQPPTTLISTVFTRAAATLQAMATTAVLTTTFQAP TDSKGSLETIPFTEISNLTLNTGNVYNPTALSMSNVESSTMNKTASWEGREASPGSSSQGSV PENQYGLPFEKWLLIGSLLFGVLFLVIGLVLLGRILSESLRRKRYSRLDYLINGIYVDI

Signal sequence.

amino acids 1-25

Transmembrane domain.

amino acids 384-405

N-glycosylation sites.

amino acids 72-76, 222-226, 251-255, 327-331, 352-356

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 415-419

Tyrosine kinase phosphorylation site.

amino acids 50-57

N-myristoylation sites.

amino acids 4-10, 48-54, 315-321

FIGURE 329

 $\tt CTCCCACGGTGTCCAGCGCCCAGA{\color{blue} \textbf{ATG}} CGGCTTCTGGTCCTGCTATGGGGTTGCCTGCTGCT$ CCCAGGTTATGAAGCCCTGGAGGGCCCAGAGGAAATCAGCGGGTTCGAAGGGGACACTGTGT CCCTGCAGTGCACCTACAGGGAAGAGCTGAGGGACCACCGGAAGTACTGGTGCAGGAAGGGT GGGATCCTCTCTCTCGCTGCTCTGGCACCATCTATGCAGAAGAAGAAGGCCAGGAGACAAT GAAGGGCAGGGTGTCCATCCGTGACAGCCGCCAGGAGCTCTCGCTCATTGTGACCCTGTGGA ACCTCACCCTGCAAGACGCTGGGGAGTACTGGTGTGGGGTCGAAAAACGGGGCCCCGATGAG TCTTTACTGATCTCTGTTCGTCTTTCCAGGACCCTGCTGTCCTCCCCTCCCCTTCTCCCAC CTTCCAGCCTCTGGCTACAACACGCCTGCAGCCCAAGGCAAAAGCTCAGCAAACCCAGCCC CAGGATTGACTTCTCCTGGGCTCTACCCGGCAGCCACCACAGCCAAGCAGGGGAAGACAGGG GCTGAGGCCCCTCCATTGCCAGGGACTTCCCAGTACGGCACGAAAGGACTTCTCAGTACAC AGGAACCTCTCCTCACCCAGCGACCTCTCCTCCTGCAGGGAGCTCCCGCCCCCCATGCAGC TGGACTCCACCTCAGCAGAGGACACCAGTCCAGCTCTCAGCAGTGGCAGCTCTAAGCCCAGG GTGTCCATCCCGATGGTCCGCATACTGGCCCCAGTCCTGGTGCTGCTGAGCCTTCTGTCAGC CGCAGGCCTGATCGCCTTCTGCAGCCACCTGCTCCTGTGGAGAAAGGAAGCTCAACAGGCCA CGGAGACACAGAGGAACGAGAAGTTCTGGCTCTCACGCTTGACTGCGGAGGAAAAGGAAGCC CCTTCCCAGGCCCCTGAGGGGGACGTGATCTCGATGCCTCCCCTCCACACATCTGAGGAGGA GCTGGGCTTCTCGAAGTTTGTCTCAGCG**TAG**GGCAGGAGGCCCTCCTGGCCAGGCCAGCAGT GAAGCAGTATGGCTGGCTGGATCAGCACCGATTCCCGAAAGCTTTCCACCTCAGCCTCAGAG TCCAGCTGCCCGGACTCCAGGGCTCTCCCCACCCTCCCCAGGCTCTCCTCTTGCATGTTCCA GCCTGACCTAGAAGCGTTTGTCAGCCCTGGAGCCCAGAGCGGTGGCCTTGCTCTTCCGGCTG GAGACTGGGACATCCCTGATAGGTTCACATCCCTGGGCAGAGTACCAGGCTGCTGACCCTCA GCAGGGCCAGACAAGGCTCAGTGGATCTGGTCTGAGTTTCAATCTGCCAGGAACTCCTGGGC TGGCGTCCTCAGACTTAGTCCCACGGTCTCCTGCATCAGCTGGTGATGAAGAGGAGCATGCT GGGGTGAGACTGGGATTCTGGCTTCTCTTTGAACCACCTGCATCCAGCCCTTCAGGAAGCCT GTGAAAAACGTGATTCCTGGCCCCACCAAGACCCACCAAAACCATCTCTGGGCTTGGTGCAG GACTCTGAATTCTAACAATGCCCAGTGACTGTCGCACTTGAGTTTGAGGGCCCAGTGGGCCTG ATGAACGCTCACACCCCTTCAGCTTAGAGTCTGCATTTGGGCTGTGACGTCTCCACCTGCCC CAATAGATCTGCTCTGCGACACCAGATCCACGTGGGGACTCCCCTGAGGCCTGCTAAG TCCAGGCCTTGGTCAGGTCAGGTGCACATTGCAGGATAAGCCCAGGACCGGCACAGAAGTGG TTGCCTTTNCCATTTGCCCTCCCTGGNCCATGCCTTCTTGCCTTTGGAAAAAATGATGAAGA AAACCTTGGCTCCTTGTCTGGAAAGGGTTACTTGCCTATGGGTTCTGGTGGCTAGAGA GAAAAGTAGAAAACCAGAGTGCACGTAGGTGTCTAACACAGAGGAGAGTAGGAACAGGGCGG ATACCTGAAGGTGACTCCGAGTCCAGCCCCCTGGAGAAGGGGTCGGGGGTGGTGAAAGTA GCACAACTACTATTTTTTTTTTTTTTCCATTATTATTGTTTTTTAAGACAGAATCTCGTGCT GCTGCCCAGGCTGGAGTGCAGTGGCACGATCTGCAAACTCCGCCTCCTGGGTTCAAGTGATT TTTGTACTTTTAGTAGAGATGGGGTTTCACCATGTTGGCCAGGCTGGTCTTGAACTCCTGAC CTCAAATGAGCCTCCTGCTTCAGTCTCCCAAATTGCCGGGATTACAGGCATGAGCCACTGTG TCTGGCCCTATTTCCTTTAAAAAGTGAAATTAAGAGTTGTTCAGTATGCAAAACTTGGAAAG ATGGAGGAGAAAAGGAAAGGAAGAAAAATGTCACCCATAGTCTCACCAGAGACTATCAT TATTTCGTTTTGTTGTACTTCCTTCCACTCTTTTCTTCTTCACATAATTTGCCGGTGTTCTT TTTACAGAGCAATTATCTTGTATACAACTTTGTATCCTGCCTTTTCCACCTTATCGTTCC GCTGCATAAAAAAAAAAAAA

FIGURE 330

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44196</pre>

<subunit 1 of 1, 332 aa, 1 stop

<MW: 36143, pI: 5.89, NX(S/T): 1

MRLLVLLWGCLLLPGYEALEGPEEISGFEGDTVSLQCTYREELRDHRKYWCRKGGILFSRCS GTIYAEEGQETMKGRVSIRDSRQELSLIVTLWNLTLQDAGEYWCGVEKRGPDESLLISLFV FPGPCCPPSPSPTFQPLATTRLQPKAKAQQTQPPGLTSPGLYPAATTAKQGKTGAEAPPLPG TSQYGHERTSQYTGTSPHPATSPPAGSSRPPMQLDSTSAEDTSPALSSGSSKPRVSIPMVRI LAPVLVLLSLLSAAGLIAFCSHLLLWRKEAQQATETQRNEKFWLSRLTAEEKEAPSQAPEGD VISMPPLHTSEEELGFSKFVSA

Important features:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 248-269

N-glycosylation site.

amino acids 96-99

Fibrinogen beta and gamma chains C-terminal domain.

amino acids 104-113

Ig like V-type domain:

amino acids 13-128